Theory of many-fermion systems

D.Belitz

Department of Physics and Materials Science Institute
University of Oregon,
Eugene, OR 97403

T.R.Kirkpatrick

Institute for Physical Science and Technology, and Department of Physics
University of Maryland,
College Park, MD 20742

(February 1, 2008)

A general field–theoretical description of many–fermion systems, both with and without quenched disorder, is developed. Starting from the basic Grassmannian action for interacting fermions, we first bosonize the theory by introducing composite matrix variables that correspond to two–fermion excitations and integrating out the fermion degrees of freedom. The saddle point solution of the resulting matrix field theory reproduces a disordered Hartree–Fock approximation, and an expansion to Gaussian order about the saddle point corresponds to a disordered RPA–like theory. In the clean limit they reduce to the ordinary Hartree–Fock and random–phase approximations. We first concentrate on disordered systems, and perform a symmetry analysis that allows for a systematic separation of the massless modes from the massive ones. By treating the massive modes in a simple approximation, one obtains a technically satisfactory derivation of the generalized nonlinear sigma–model that has been used in the theory of metal–insulator transitions. The theory also allows for the treatment of other phase transitions in the disordered Fermi liquid. We further use renormalization group techniques to establish the existence of a disordered Fermi–liquid fixed point, and show that it is stable for all dimensions $d > 2$. The so–called weak–localization effects can be understood as corrections to scaling near this fixed point. The general theory also allows for studying the clean limit. For that case we develop a loop expansion that corresponds to an expansion in powers of the screened Coulomb interaction, and that represents a systematic improvement over RPA. We corroborate the existence of a Fermi–liquid fixed point that is stable for all dimensions $d > 1$, in agreement with other recent field–theoretical treatments of clean Fermi liquids.

Contents

I. INTRODUCTION

The many–fermion problem has a long history, due to its importance with respect to electrons in condensed matter systems.[1–3] In recent years there has been a renewed interest in fundamental aspects of this problem, especially for electrons at low or zero temperature, both with and without quenched disorder. The important physical systems for which these studies are relevant include, superconductors, both high–T_c and conventional ones, doped semiconductors in both their metallic and insulating phases, amorphous alloys, and Quantum Hall systems.

Historically, there have been two important techniques to tackle the many–fermion problem: Landau's phenomenological Fermi–liquid theory,[4] and the microscopic many–body perturbation theory or Feynman diagram approach,[2,3] which starts from a canonically quantized Hamiltonian of the problem. While both techniques are very useful, and have yielded many important results, either one of them has serious limitations. Fermi–liquid theory has been extended to include disorder,[5,6] but often a microscopic approach is preferred. For the microscopic many–body perturbation theory, the inclusion of disorder is extremely awkward, mostly because of a large number of diagrammatic contributions that are individually finite, but that ultimately cancel each other. More recent approaches have used functional or field–theoretical methods,[7] which allow for the application of renormalization group (RG) techniques. A significant recent advance in this area has been Shankar's RG technique[8] for clean fermion systems that is applied directly to the Grassmannian field theory for fermions. This approach, which is still in its infancy, is very promising. For instance, it has already led to a derivation of Fermi–liquid theory from a microscopic starting point.[8,9] It also has been used to provide RG derivations of the Cooper instability problem, and of RPA–like screening.[8] Other applications, for instance a discussion of corrections to scaling near the RG fixed point (FP) that describes the clean Fermi liquid, should be possible. However, for reasons that will be discussed in detail in the present paper, it is not easy to include the effects of quenched disorder into this approach either.

An alternative approach for clean electronic systems has been pursued by Houghton, Marston, and others,[10] and by Fröhlich and Götschmann.[11] These authors have generalized the bosonization techniques that have been used successfully in $d = 1$ to higher dimensions. In Ref. 11 the bosonization was also combined with RG techniques. These theories have been used, for example, to rederive certain nonanalyticities that occur in Fermi–liquid theory, and to study the possibility of marginal Fermi liquids in dimensions $d \geq 2$. Finally, clean electronic systems have been bosonized by means of a Hubbard–Stratonovich transformation and the use of the classical Hubbard–Stratonovich field as the fundamental field of the theory.[12] Within this latter approach, disorder has been included in the single-particle Green function at the level of the lowest order Born approximation.[13] More sophisticated effects of quenched disorder have so far not been included into any of these approaches.

For systems of disordered electrons, a completely different theory has been in use for some time, that describes composite fermionic, i.e. effectively bosonic, degrees of freedom. For noninteracting electrons, this effective field theory takes the form of a nonlinear sigma–model,[14] which was generalized to the case of interacting electrons by Finkel'stein.[15,16] It is custom tailored for the description of the metal–insulator transition near $d = 2$, and it does not allow for the clean limit to be taken. No technically satisfactory derivation of the interacting theory has ever been given, but rather its structure has been guessed, based on more rigorous derivations of Wegner's effective model for the metal–insulator transition in noninteracting electronic systems.[17,18] The key idea underlying these effective field theories is to keep explicitly only those degrees of freedom that are likely to be relevant for the problem under consideration, and to integrate out all others in some simple approximation.

It is clearly desirable to develop a more flexible field–theoretical approach for the many–fermion problem. Some of the desired features of such a theory are as follows. (1) Both clean and disordered systems should be describable within a single framework. (2) The theory should allow for a RG description of both clean and disordered Fermi liquids, of the metal–insulator transition, as well as of other quantum phase transitions, like e.g. the recently investigated magnetic ones.[19–21] (3) It should offer flexibility with respect to how many, and which, degrees of freedom one keeps explicitly, depending on the problem under consideration. (4) It should provide a satisfactory derivation of the nonlinear sigma–model used to describe the metal–insulator transition in interacting disordered electronic systems. (5) It should allow for explicit calculations of physically relevant observables, such as thermodynamic and transport properties. While all of the existing theories fulfill some of these criteria, none of them meet all of them simultaneously.

This is the first of two papers where we develop such a theory. In the present paper we discuss our general physical ideas, and mainly focus on the disordered interacting fermion problem, although we will discuss some aspects of the clean limit. In a future, second paper we plan to use the same approach to thoroughly discuss clean interacting fermion systems, as well as the connection between our approach and others. The outline of this paper is as follows. In Sec. II we introduce our starting point, a microscopic model for an electron fluid, in general in the presence of quenched disorder. We then show how to transform this Grassmannian field theory

into one for classical or bosonic fields, and discuss the physical motivation for this transformation. In this section we also construct a saddle–point solution for the resulting composite–fermion field theory, and expand to Gaussian order about the saddle point. The result is equivalent to what one obtains within many–body diagrammatic theory from the random phase approximation (RPA), modified by disorder. In Sec. III we perform a symmetry analysis of our field theory. The results, together with some conclusions drawn from Sec. II, suggest slightly separate ways to proceed from here for clean and disordered fermions, respectively. Focusing on the latter, we use the symmetry analysis to identify and classify all of the slow or soft modes of the system. This is crucial for a RG description of the problem. Using these results, we then derive an effective field theory for the slow modes in a disordered fermion system. We identify a FP that describes a disordered Fermi liquid, and show that it is stable for $d > 2$. We then show that the so–called weak–localization effects can be interpreted as corrections to scaling near this FP. We derive Finkel'stein's generalization of Wegner's nonlinear sigma–model that has been used to describe metal–insulator transitions near $d = 2$, and discuss the critical FP that corresponds to this quantum phase transition. The section is concluded by a brief discussion of a magnetic phase transition that is described by the theory, and of the relation between the theory presented here and earlier theories of magnetic transitions in disordered systems. In Sec. IV we consider the clean limit of our theory and discuss the clean Fermi–liquid FP. There we show that certain nonanalyticities that appear in Fermi–liquid theory can be interpreted as corrections to scaling near this FP, and we point out far–reaching analogies between clean and disordered Fermi systems. In Sec. V we present our conclusions. Appendix A spells out a technical point that one encounters in Sec. II, and Appendix B contains a pedagogical discussion of an $O(N)$ symmetric ϕ^4 theory.

II. MATRIX FIELD THEORY

A. Grassmannian field theory

Let us start with a field–theoretical description of a system of interacting, disordered fermions in terms of Grassmann variables.[7] The partition function of the system is

$$Z = \int D[\bar\psi, \psi] \ e^{S[\bar\psi, \psi]} \quad . \tag{2.1}$$

Here the functional integration is with respect to Grassmann valued fields, $\bar\psi$ and ψ, and the action, S, is given by

$$S = -\int dx \sum_\sigma \bar\psi_\sigma(x)\, \partial_\tau\, \psi_\sigma(x) + S_0 + S_{dis} + S_{int} \quad . \tag{2.2a}$$

We use a $(d+1)$-vector notation, with $x = (\mathbf{x}, \tau)$, and $\int dx = \int_V d\mathbf{x} \int_0^\beta d\tau$. $\mathbf{x}$ denotes position, τ imaginary time, V is the system volume, $\beta = 1/T$ is the inverse temperature, σ is the spin label, and we use units such that $\hbar = k_B = 1$. S_0 describes free fermions with chemical potential μ,

$$S_0 = \int dx \sum_\sigma \ \bar\psi_\sigma(x) \left(\frac{\nabla^2}{2m} + \mu \right) \psi_\sigma(x) \quad , \tag{2.2b}$$

with m the fermion mass. The Laplacian will be denoted by ∇^2 throughout. We will mostly be concerned with systems at $T = 0$, where $\mu = \epsilon_F$, with $\epsilon_F = k_F^2/2m$ the Fermi energy, and k_F the Fermi momentum. S_{dis} describes a static random potential, $u(\mathbf{x})$, coupling to the fermionic number density,

$$S_{dis} = -\int dx \sum_\sigma \ u(\mathbf{x})\, \bar\psi_\sigma(x)\, \psi_\sigma(x) \quad , \tag{2.2c}$$

and S_{int} describes a spin–independent two–particle interaction,

$$S_{int} = -\frac{1}{2} \int dx_1\, dx_2 \sum_{\sigma_1,\sigma_2} v(\mathbf{x}_1 - \mathbf{x}_2)$$
$$\times \bar\psi_{\sigma_1}(x_1)\, \bar\psi_{\sigma_2}(x_2)\, \psi_{\sigma_2}(x_2)\, \psi_{\sigma_1}(x_1) \quad . \tag{2.2d}$$

The interaction potential $v(\mathbf{x})$ will be specified below.

For simplicity, the random potential in Eq. (2.2c) is taken to be Gaussian distributed with a second moment that is given by a function $U(\mathbf{x})$ with the dimension of an energy density,

$$\{u(\mathbf{x})\, u(\mathbf{y})\}_{dis} = \frac{1}{\pi N_F}\, U(\mathbf{x} - \mathbf{y}) \quad , \tag{2.3}$$

where $\{\dots\}_{dis}$ denotes the disorder average. N_F is a normalization factor with the dimension of a density of states. One could use the free electron density of states at the Fermi level, but it will turn out to be more convenient to include some trivial disorder and interaction renormalizations. We will explicitly define N_F in Eq. (2.46) below. The full physical density of states at the Fermi level will also be encountered later, and will be denoted by $N(\epsilon_F)$. More general disorder potentials can be considered, but as long as they couple only to the electron number density, the difference between the more general potentials and Eq. (2.3) can be shown to be RG irrelevant at the disordered Fermi–liquid FP. For other purposes, e.g. a description of the Anderson–Mott metal–insulator transition in the system, the difference is likewise expected to be RG irrelevant.

Since the system contains quenched disorder, it is necessary to average the free energy or $\ln Z$. This is accomplished by means of the replica trick,[22] which is based on the identity

$$\ln Z = \lim_{n\to 0} (Z^n - 1)/n \quad . \tag{2.4}$$

Introducing n identical replicas of the system (with n an integer), labeled by the index α, and carrying out the disorder average, we obtain

$$\tilde{Z} \equiv \{Z^n\}_{dis} = \int \prod_{\alpha=1}^{n} D\left[\bar{\psi}^\alpha, \psi^\alpha\right] \exp[\tilde{S}] \quad . \qquad (2.5)$$

We again separate $\tilde{S}$ into free, disordered and interaction parts as follows,

$$\tilde{S} = \sum_{\alpha=1}^{n} \left(\tilde{S}_0^\alpha + \tilde{S}_{dis}^\alpha + \tilde{S}_{int}^\alpha\right) \quad . \qquad (2.6a)$$

It is useful to go to a Fourier representation with wave vectors $\mathbf{k}$ and fermionic Matsubara frequencies $\omega_n = 2\pi T(n+1/2)$, and a $(d+1)$-vector notation, $k = (\mathbf{k}, \omega_n)$. Then

$$\tilde{S}_0^\alpha = \sum_{k,\sigma} \bar{\psi}_\sigma^\alpha(k) \left[i\omega_n - \mathbf{k}^2/2m + \mu\right] \psi_\sigma^\alpha(k) \quad , \qquad (2.6b)$$

and,

$$\begin{aligned}
\tilde{S}_{dis}^\alpha = \frac{1}{2\pi N_F} &\sum_{\beta=1}^{n} \sum_{\{\mathbf{k}_i\}} \sum_{n,m} \sum_{\sigma,\sigma'} \delta_{\mathbf{k}_1+\mathbf{k}_3,\mathbf{k}_2+\mathbf{k}_4} \\
&\times U(\mathbf{k}_1 - \mathbf{k}_2)\, \bar{\psi}_{n\sigma}^\alpha(\mathbf{k}_1)\, \psi_{n\sigma}^\alpha(\mathbf{k}_2)\, \bar{\psi}_{m\sigma'}^\beta(\mathbf{k}_3) \\
&\qquad \times \psi_{m\sigma'}^\beta(\mathbf{k}_4) \quad , \qquad (2.6c)
\end{aligned}$$

$$\begin{aligned}
\tilde{S}_{int}^\alpha = -\frac{T}{2} &\sum_{\sigma_1,\sigma_2} \sum_{\{k_i\}} \delta_{k_1+k_2,k_3+k_4}\, v(\mathbf{k}_2 - \mathbf{k}_3) \\
&\times \bar{\psi}_{\sigma_1}^\alpha(k_1)\, \bar{\psi}_{\sigma_2}^\alpha(k_2)\, \psi_{\sigma_2}^\alpha(k_3)\, \psi_{\sigma_1}^\alpha(k_4) \quad . \qquad (2.6d)
\end{aligned}$$

Here we have defined

$$\psi_{n\sigma}(\mathbf{k}) \equiv \psi_\sigma(k) = \sqrt{T/V} \int dx\, e^{-i(\mathbf{k}\mathbf{x}-\omega_n\tau)}\, \psi_\sigma(x) \quad , \tag{2.7a}$$

$$\bar{\psi}_{n\sigma}(\mathbf{k}) \equiv \bar{\psi}_\sigma(k) = \sqrt{T/V} \int dx\, e^{i(\mathbf{k}\mathbf{x}-\omega_n\tau)}\bar{\psi}_\sigma(x) \quad . \tag{2.7b}$$

Ultimately, we will be interested in long–wavelength, low–frequency processes. For clean electrons, this in general means that only the scattering of electrons and holes close to the Fermi surface is important. In a clean Fermi liquid this is true because the lifetimes of the single–particle excitations become infinitely long as the Fermi surface is approached. For the disordered electron problem the situation is different. All single–particle momentum and frequency excitations are damped by the single–particle relaxation rate, $1/\tau_{rel}$, which in general is not small. For this case, the only slow or soft excitations are two–particle excitations. This implies that the soft
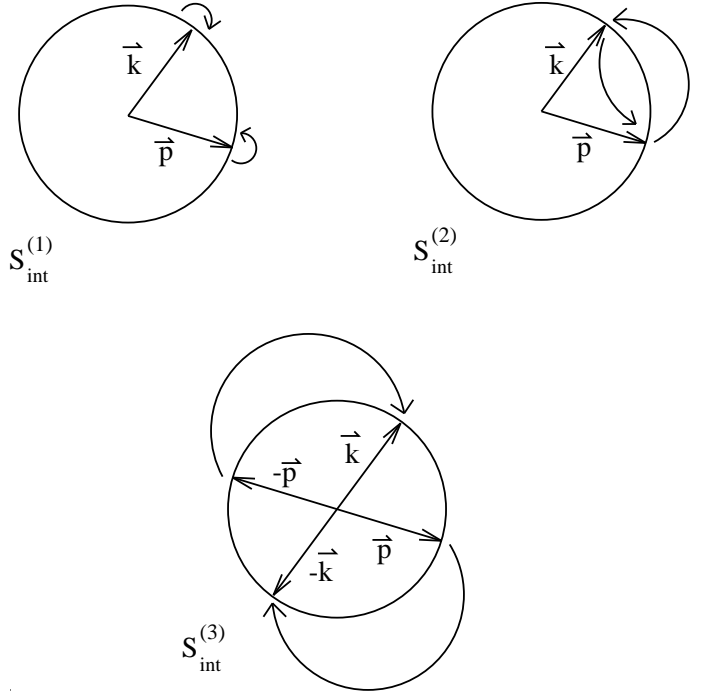

FIG. 1. Typical small–angle (1), large–angle (2), and $2k_F$-scattering processes (3) near the Fermi surface in $d = 2$.

excitations in the two cases are fundamentally different, and that clean systems have in general more soft modes than disordered ones. This observation will be of great importance below, and will suggest somewhat different effective field theories for dealing with clean and disordered systems, respectively. In fact, this difference is the reason why the approaches of Shankar[8] and Houghton and Marston[10] cannot be simply extended to the dirty case.

To examine this important point in more detail, we first consider the clean case. It is customary and convenient to divide the possible scattering processes into three classes: (1) small–angle scattering, (2) large–angle scattering, and (3) $2k_F$-scattering. These classes are also referred to as the particle–hole channel for classes (1) and (2), and the particle–particle or Cooper channel for class (3), respectively. The corresponding scattering processes are schematically depicted in Fig. 1. We next make explicit the phase space decomposition that is inherent to this classification by writing $\tilde{S}_{int}^\alpha$, Eq. (2.6d), as

$$\tilde{S}_{int}^\alpha = \tilde{S}_{int}^{\alpha\,(1)} + \tilde{S}_{int}^{\alpha\,(2)} + \tilde{S}_{int}^{\alpha\,(3)} \quad , \qquad (2.8a)$$

where,

$$\begin{aligned}
\tilde{S}_{int}^{\alpha\,(1)} = -\frac{T}{2} &\sum_{\sigma_1\sigma_2} \sum_{k,p} {\sum_q}'\, v(\mathbf{q})\, \bar{\psi}_{\sigma_1}^\alpha(k)\, \bar{\psi}_{\sigma_2}^\alpha(p+q) \\
&\times \psi_{\sigma_2}^\alpha(p)\psi_{\sigma_1}^\alpha(k+q) \quad , \qquad (2.8b)
\end{aligned}$$

$$\tilde{S}_{int}^{\alpha\,(2)} = -\frac{T}{2} \sum_{\sigma_1\sigma_2} \sum_{k,p} {\sum_q}'\, v(\mathbf{p} - \mathbf{k})\, \bar{\psi}_{\sigma_1}^\alpha(k)\, \bar{\psi}_{\sigma_2}^\alpha(p+q)$$

$$\times \psi^\alpha_{\sigma_2}(k+q)\psi^\alpha_{\sigma_1}(p) \quad , \qquad (2.8c)$$

$$\tilde{S}^{\alpha\,(3)}_{int} = -\frac{T}{2}\sum_{\sigma_1\neq\sigma_2}\sum_{k,p}\sum_q{}' v(\mathbf{k}+\mathbf{p})\,\bar\psi^\alpha_{\sigma_1}(-k)\,\bar\psi^\alpha_{\sigma_2}(k+q)$$

$$\times \psi^\alpha_{\sigma_2}(-p+q)\,\psi^\alpha_{\sigma_1}(p) \quad . \qquad (2.8d)$$

Here the prime on the q-summation indicates that only momenta up to some momentum cutoff, λ, are integrated over. This restriction is necessary to avoid double counting, since each of the three expressions, Eqs. (2.8b) - (2.8d), represent all of $\tilde{S}^\alpha_{int}$ if all wave vectors are summed over. In general the long–wavelength physics we are interested in will not depend on λ. Finally, one often writes the Eqs. (2.8) in terms of singlet (s) and triplet (t) contributions. We introduce spinors

$$\psi^\alpha_n(\mathbf{x}) = \begin{pmatrix}\psi^\alpha_{n\uparrow}(\mathbf{x})\\ \psi^\alpha_{n\downarrow}(\mathbf{x})\end{pmatrix} \quad , \qquad (2.9a)$$

and their Fourier transforms

$$\psi^\alpha(k) \equiv \psi^\alpha_n(\mathbf{k}) = \begin{pmatrix}\psi^\alpha_{n\uparrow}(\mathbf{k})\\ \psi^\alpha_{n\downarrow}(\mathbf{k})\end{pmatrix} \quad , \qquad (2.9b)$$

as well as their adjoints, $\bar\psi^\alpha(k)$, and a scalar product in spinor space, $(\psi,\psi) = \bar\psi\cdot\psi$, where the dot denotes the matrix product. Then we can write the interaction term as

$$\tilde{S}^\alpha_{int} = \tilde{S}^{\alpha\,(s)}_{int} + \tilde{S}^{\alpha\,(t)}_{int} + \tilde{S}^{\alpha\,(3)}_{int} \quad , \qquad (2.10a)$$

with,

$$\tilde{S}^{\alpha\,(s)}_{int} = -\frac{T}{2}\sum_{k,p}\sum_q{}' \Gamma^{(s)}_{k,p}(q)\,\big(\psi^\alpha(k),s_0\psi^\alpha(k+q)\big)$$

$$\times\big(\psi^\alpha(p+q),s_0\psi^\alpha(p)\big) \quad , \qquad (2.10b)$$

$$\tilde{S}^{\alpha\,(t)}_{int} = -\frac{T}{2}\sum_{k,p}\sum_q{}' \Gamma^{(t)}_{k,p}(q)\sum_{i=1}^3\big(\psi^\alpha(k),s_i\psi^\alpha(k+q)\big)$$

$$\times\big(\psi^\alpha(p+q),s_i\psi^\alpha(p)\big) \quad . \qquad (2.10c)$$

Here $s_j = i\sigma_j$, with $\sigma_{1,2,3}$ the Pauli matrices, and $s_0 = \sigma_0$ is the 2×2 identity matrix. We have also defined the singlet (s) and triplet (t) interaction amplitudes

$$\Gamma^{(t)}_{k,p}(q) = \frac{1}{2}\,v(\mathbf{p}-\mathbf{k}) \quad , \qquad (2.11a)$$

and

$$\Gamma^{(s)}_{k,p}(q) = v(\mathbf{q}) - \frac{1}{2}\,v(\mathbf{p}-\mathbf{k}) \quad . \qquad (2.11b)$$

In addition we define the Cooper channel or $2k_F$-scattering amplitude

$$\Gamma^{(c)}_{k,p}(q) = v(\mathbf{k}+\mathbf{p}) \quad . \qquad (2.11c)$$

The effective interaction potentials that appear in Eqs. (2.10) and (2.11) are all given by the basic interaction potential v, taken at different momenta. $\tilde{S}^{\alpha\,(1)}_{int}$, and $\tilde{S}^{\alpha\,(s)}_{int}$ as well, contains the direct scattering contribution, or $v(\mathbf{q})$, with $\mathbf{q}$ the restricted ($|\mathbf{q}| < \lambda << k_F$) momentum. If v is chosen to be a bare Coulomb interaction, then the direct scattering contribution leads to singularities in perturbation theory that indicate the need for infinite resummations to incorporate screening. For simplicity, we assume that this procedure has already been carried out, and take v to be a statically screened Coulomb interaction. For small $|\mathbf{q}|$, it is then sufficient to replace $v(\mathbf{q})$ in Eq. (2.11b) by a number. The moduli of the other momenta, $\mathbf{k}$ and $\mathbf{p}$ in Eqs. (2.10), are equal to k_F for the most important scattering processes since, as mentioned above, the soft or slow excitations in the clean Fermi system involve particles and holes close to the Fermi surface. The angular dependences of these coupling constants is usually taken into account by expanding them in Legendre polynomials on the Fermi surface.[4] For future reference we note that if only the zero angular momentum channel, $\ell = 0$, is retained, then Eqs. (2.10) and (2.11) are valid with $\Gamma^{(t)}_{k,p}(0)$, $\Gamma^{(s)}_{k,p}(0)$, and $\Gamma^{(c)}_{k,p}(0)$ replaced by numbers, $\Gamma^{(s)}$, $\Gamma^{(t)}$, and $\Gamma^{(c)}$, which are the Fermi surface averages of the respective interaction amplitudes. Shankar,[8] in his RG approach, has explicitly derived these results, and has shown that all corrections to these approximations are RG irrelevant with respect to a RG fixed point that describes a Fermi liquid.

For disordered Fermi systems the situation is different, since single–particle excitations are massive modes with a mass that is proportional to the inverse elastic mean–free time. In Section III we will give a detailed discussion of the soft or slow modes in a disordered Fermi system. The conclusion will be that, of the modes that appear in $\tilde{S}_{int}$, the dominant soft modes are those that involve fluctuations of either the particle number density, n_n, or the spin density, $\mathbf{n}_s$, in the particle–hole channel, or density fluctuations, n_c, in the particle–particle channel.[23] In terms of our spinors, the latter are given as (cf. Eqs. (2.10))

$$n^\alpha_n(q) = \sqrt{T/V}\sum_k\big(\psi^\alpha(k),s_0\psi^\alpha(k+q)\big) \quad , \qquad (2.12a)$$

$$\mathbf{n}^\alpha_s(q) = \sqrt{T/V}\sum_k\big(\psi^\alpha(k),\mathbf{s}\psi^\alpha(k+q)\big) \quad , \qquad (2.12b)$$

$$n^\alpha_c(q) = \sqrt{T/V}\sum_k\bar\psi^\alpha_\uparrow(k)\,\bar\psi^\alpha_\downarrow(-k+q) \quad . \qquad (2.12c)$$

In Eqs. (2.10) this implies that we should project onto these density fluctuations. This projection is performed in Appendix A. The net result is that the interaction amplitudes can be replaced by constants, namely their angular averages with a distribution of momenta near

the Fermi surface. We thus obtain the interaction part of the action in the form,

$$\tilde{S}_{int}^{\alpha\,(s)} = -\frac{T}{2}\,\Gamma^{(s)} \sum_{k,p}\sum_{q}{}' \left(\psi^\alpha(k), s_0\psi^\alpha(k+q)\right)$$
$$\times \left(\psi^\alpha(p+q), s_0\psi^\alpha(p)\right) \quad , \qquad (2.13a)$$

$$\tilde{S}_{int}^{\alpha\,(t)} = -\frac{T}{2}\,\Gamma^{(t)} \sum_{k,p}\sum_{q}{}'\sum_{i=1}^{3} \left(\psi^\alpha(k), s_i\psi^\alpha(k+q)\right)$$
$$\times \left(\psi^\alpha(p+q), s_i\psi^\alpha(p)\right) \quad . \qquad (2.13b)$$

$$\tilde{S}_{int}^{\alpha\,(3)} = -\frac{T}{2}\,\Gamma^{(c)} \sum_{\sigma_1\neq\sigma_2}\sum_{k,p}\sum_{q}{}' \bar{\psi}_{\sigma_1}^\alpha(k)\,\bar{\psi}_{\sigma_2}^\alpha(-k+q)$$
$$\times \psi_{\sigma_2}^\alpha(p+q)\,\psi_{\sigma_1}^\alpha(-p) \quad . \qquad (2.13c)$$

The disordered Fermi system is thus described by only three interaction parameters, $\Gamma^{(s)}$, $\Gamma^{(t)}$, and $\Gamma^{(c)}$, which are given by Eq. (A4b) and analogous expressions. As long as we are only interested in physics that is dominated by soft modes, this restriction is exact. Clean Fermi systems, on the other hand, in general require an infinite number of Fermi–liquid parameters, and the restriction to the Eqs. (2.13), with the three interaction constants related to the Fermi–liquid parameters in the $\ell = 0$ channel, constitutes an approximation. Physically this difference is due to the fact that there are more soft modes in the clean case than in the dirty case.

B. Composite variables: Matrix field theory

1. The action in terms of Q-matrices

As noted above, the slow modes for the disordered fermion problem are fluctuations of products of fermion operators. It is therefore convenient to transform to a field theory in terms of these composite variables. For both technical and physical reasons, which will become clear in Sec. III, it is convenient to go to a bispinor and, eventually, to a spin–quaternion representation. We define a bispinor[24]

$$\eta_n^\alpha(\mathbf{x}) = \frac{1}{\sqrt{2}}\left(\begin{array}{c} \bar{\psi}_n^\alpha(\mathbf{x}) \\ s_2\,\psi_n^\alpha(\mathbf{x}) \end{array}\right) = \frac{1}{\sqrt{2}}\left(\begin{array}{c} \bar{\psi}_{n\uparrow}^\alpha(\mathbf{x}) \\ \bar{\psi}_{n\downarrow}^\alpha(\mathbf{x}) \\ \psi_{n\downarrow}^\alpha(\mathbf{x}) \\ -\psi_{n\uparrow}^\alpha(\mathbf{x}) \end{array}\right) \quad , \qquad (2.14a)$$

and an adjoint bispinor[25]

$$(\eta^+)_n^\alpha(\mathbf{x}) = ic\eta_n^\alpha(\mathbf{x}) = \frac{i}{\sqrt{2}}\left(\begin{array}{c} -\psi_{n\uparrow}^\alpha(\mathbf{x}) \\ -\psi_{n\downarrow}^\alpha(\mathbf{x}) \\ \bar{\psi}_{n\downarrow}^\alpha(\mathbf{x}) \\ -\bar{\psi}_{n\uparrow}^\alpha(\mathbf{x}) \end{array}\right) \quad , \qquad (2.14b)$$

with c the charge–conjugation matrix

$$c = \left(\begin{array}{cc} 0 & s_2 \\ s_2 & 0 \end{array}\right) = i\tau_1 \otimes s_2 \quad . \qquad (2.14c)$$

The four degrees of freedom represented by the bispinor are the particle–hole or number density degrees of freedom, and the two spin degrees of freedom. We have also defined a basis in spin–quaternion space as $\tau_r \otimes s_i$ $(r, i = 0, 1, 2, 3)$, with $\tau_0 = s_0$ the 2×2 identity matrix, and $\tau_j = -s_j = -i\sigma_j$ $(j = 1, 2, 3)$, with σ_j the Pauli matrices. In this basis, the channels $r = 0, 3$ and $r = 1, 2$ describe the particle–hole and particle–particle degrees of freedom, and the channels $i = 0$ and $i = 1, 2, 3$ describe the spin–singlet and spin–triplet, respectively.

It is convenient to define a scalar product in bispinor space as

$$\left(\eta_n^\alpha, \eta_m^\beta\right) = \sum_{j=1}^{4} {}_j\eta_n^\alpha\,{}^j\eta_m^\beta \quad , \qquad (2.15)$$

where ${}^j\eta$ denotes the elements of η, and ${}_j\eta$ denotes the elements of η^+. The adjoint of any operator A in bispinor space with respect to this scalar product is given by $c^T A^T c$. In terms of the bispinors, the terms on the right–hand side of Eq. (2.6a) can be written

$$\tilde{S}_0 = -i\sum_{\alpha,k}\left(\eta^\alpha(k), [i\omega_n - \mathbf{k}^2/2m + \mu]\,\eta^\alpha(k)\right) \quad , \qquad (2.16a)$$

$$\tilde{S}_{dis} = \frac{-1}{2\pi N_F}\sum_{\alpha,\beta}\sum_{n,m}\sum_{\{\mathbf{k}_i\}} \delta_{\mathbf{k}_1+\mathbf{k}_3,\mathbf{k}_2+\mathbf{k}_4}\,U(\mathbf{k}_1 - \mathbf{k}_2)$$
$$\times \left(\eta_n^\alpha(\mathbf{k}_1), \eta_n^\alpha(\mathbf{k}_2)\right)\left(\eta_m^\beta(\mathbf{k}_3), \eta_m^\beta(\mathbf{k}_4)\right) \quad , \qquad (2.16b)$$

and

$$\tilde{S}_{int} = \tilde{S}_{int}^{(s)} + \tilde{S}_{int}^{(t)} + \tilde{S}_{int}^{(c)} \quad , \qquad (2.17a)$$

$$\tilde{S}_{int}^{(s)} = \frac{T\Gamma^{(s)}}{2}\sum_{\alpha}\sum_{k,p}\sum_{q}{}'\sum_{r=0,3}(-1)^r$$
$$\times \left(\eta^\alpha(k), (\tau_r \otimes s_0)\eta^\alpha(k+q)\right)$$
$$\times \left(\eta^\alpha(p+q), (\tau_r \otimes s_0)\eta^\alpha(p)\right) \quad , \qquad (2.17b)$$

$$\tilde{S}_{int}^{(t)} = \frac{T\Gamma^{(t)}}{2}\sum_{\alpha}\sum_{k,p}\sum_{q}{}'\sum_{r=0,3}(-1)^r\sum_{i=1}^{3}$$
$$\times \left(\eta^\alpha(k), (\tau_r \otimes s_i)\eta^\alpha(k+q)\right)$$
$$\times \left((\eta^\alpha(p+q), (\tau_r \otimes s_i)\eta^\alpha(p)\right) \quad , \qquad (2.17c)$$

$$\tilde{S}_{int}^{(c)} = \frac{T\Gamma^{(c)}}{2}\sum_{\alpha}\sum_{\mathbf{k},\mathbf{p}}\sum_{\mathbf{q}}{}'\sum_{n_1,n_2,m}\sum_{r=1,2}$$
$$\times \left(\eta_{n_1}^\alpha(-\mathbf{k}), (\tau_r \otimes s_0)\,\eta_{-n_1+m}^\alpha(-\mathbf{k}+\mathbf{q})\right)$$
$$\times \left(\eta_{-n_2}^\alpha(-\mathbf{p}), (\tau_r \otimes s_0)\,\eta_{n_2+m}^\alpha(-\mathbf{p}-\mathbf{q})\right) \quad . \qquad (2.17d)$$

The same phase space decomposition used to break up S_{int} can also be used to rewrite $\tilde{S}_{dis}$ as

$$\tilde{S}_{dis} = \tilde{S}_{dis}^{(1)} + \tilde{S}_{dis}^{(2)} \quad , \tag{2.18a}$$

with,

$$\tilde{S}_{dis}^{(1)} = \frac{-1}{2\pi N_F \tau_1} \sum_{\alpha,\beta} \sum_{n,m} \sum_{\mathbf{k},\mathbf{p}} {\sum_{\mathbf{q}}}' \left(\eta_n^\alpha(\mathbf{k}), \eta_n^\alpha(\mathbf{k}+\mathbf{q}) \right)$$
$$\times \left(\eta_m^\beta(\mathbf{p}), \eta_m^\beta(\mathbf{p}-\mathbf{q}) \right) \quad , \tag{2.18b}$$

$$\tilde{S}_{dis}^{(2)} = \frac{-1}{\pi N_F \tau_{rel}} \sum_{\alpha,\beta} \sum_{n,m} \sum_{\mathbf{k},\mathbf{p}} {\sum_{\mathbf{q}}}' \left(\eta_n^\alpha(\mathbf{k}), \eta_n^\alpha(\mathbf{p}) \right)$$
$$\times \left(\eta_m^\beta(\mathbf{p}+\mathbf{q}), \eta_m^\beta(\mathbf{k}+\mathbf{q}) \right) \quad . \tag{2.18c}$$

Here we have introduced two different scattering times, τ_1 and τ_{rel}, that will in general renormalize differently. The reason for their appearance is related to the potential correlation function $U(\mathbf{k}_1 - \mathbf{k}_2)$ in Eq. (2.16b), which leads to a $U(\mathbf{q})$ in $\tilde{S}_{dis}^{(1)}$, and a $U(\mathbf{k}-\mathbf{p})$ in $\tilde{S}_{dis}^{(2)}$. Projection onto the density, as in the interaction terms, then leads to Eqs. (2.18), with $1/\tau_1 = U(\mathbf{q}=0)$, and $1/\tau_{rel}$ a weighted average over $U(\mathbf{k}-\mathbf{p})$. τ_{rel} is the single–particle relaxation or scattering time.

Next we define a Grassmannian matrix field B_{12}, with $1 \equiv (n_1, \alpha_1)$, etc, by

$$B_{12}(\mathbf{x}) = \eta_1^+(\mathbf{x}) \otimes \eta_2(\mathbf{x}) \quad , \tag{2.19a}$$

or, in Fourier space and with all components written out,

$$_i^j B_{nm}^{\alpha\beta}(\mathbf{q}) = \sum_{\mathbf{k}} {}_i\eta_n^\alpha(\mathbf{k}) \, {}^j\eta_m^\beta(\mathbf{k}+\mathbf{q}) \quad , \tag{2.19b}$$

The subset of the Grassmann algebra that is given by the bilinear fields B is isomorphic to a set of classical, i.e. c-number valued, matrix fields. We exploit this isomorphism, which in particular maps the adjoint operation in ψ-space that we denoted by an overbar onto the complex conjugation operation, by introducing a classical matrix field Q, and constraining B to Q by means of a functional delta function.[26] We then use a functional integral representation of the latter, with an auxiliary or ghost field $\tilde{\Lambda}$ that plays the role of a Lagrange multiplier, and integrate out the fermions. This way we obtain,

$$\tilde{Z} = \int D[\bar{\psi}, \psi] \, e^{\tilde{S}[\bar{\psi},\psi]} \int D[Q] \, \delta[Q - B]$$
$$= \int D[\eta] \, e^{\tilde{S}[\eta]} \int D[Q] \, D[\tilde{\Lambda}] \, \exp\left(\text{tr} \left[\tilde{\Lambda}(Q - B) \right] \right)$$
$$= \int D[Q] \, D[\tilde{\Lambda}] \, e^{\mathcal{A}[Q,\tilde{\Lambda}]} \quad . \tag{2.20}$$

In Eq. (2.20) we have defined an effective action,

$$\mathcal{A}[Q, \tilde{\Lambda}] = \mathcal{A}_{dis}[Q] + \mathcal{A}_{int}[Q] + \frac{1}{2} \, \text{Tr} \ln \left(G_0^{-1} - i\tilde{\Lambda} \right)$$
$$+ \int d\mathbf{x} \, \text{tr} \left(\tilde{\Lambda}(\mathbf{x}) \, Q(\mathbf{x}) \right) \quad . \tag{2.21}$$

Here and in what follows, tr is a trace over all discrete degrees of freedom that are not shown explicitly, while Tr is a trace over all degrees of freedom, including an integral over $\mathbf{x}$.

$$G_0^{-1} = -\partial_\tau + \nabla^2/2m + \mu \quad , \tag{2.22}$$

is the inverse of the free electron Green operator, and it is clear from the structure of the Tr ln-term in Eq. (2.21) that the physical interpretation of the ghost field is that of a self–energy. In writing Eq. (2.21), we have replaced the dummy integration variables $\tilde{\Lambda}$ and Q by their transposeds, which carries a Jacobian of unity, and have anticipated the fact that $\mathcal{A}_{dis}[Q^T] = \mathcal{A}_{dis}[Q]$, and the same for $\mathcal{A}_{int}$. The latter two contributions to the action read

$$\mathcal{A}_{dis}[Q] = \mathcal{A}_{dis}^{(1)}[Q] + \mathcal{A}_{dis}^{(2)}[Q] \quad , \tag{2.23a}$$

$$\mathcal{A}_{dis}^{(1)}[Q] = \frac{-1}{2\pi N_F \tau_1} \int d\mathbf{x} \left(\text{tr} \, Q(\mathbf{x}) \right)^2 \quad , \tag{2.23b}$$

$$\mathcal{A}_{dis}^{(2)}[Q] = \frac{1}{\pi N_F \tau_{rel}} \int d\mathbf{x} \, \text{tr} \left(Q(\mathbf{x}) \right)^2 \quad , \tag{2.23c}$$

and

$$\mathcal{A}_{int}[Q] = \mathcal{A}_{int}^{(s)} + \mathcal{A}_{int}^{(t)} + \mathcal{A}_{int}^{(c)} \quad , \tag{2.24a}$$

$$\mathcal{A}_{int}^{(s)} = \frac{T\Gamma^{(s)}}{2} \int d\mathbf{x} \sum_{r=0,3} (-1)^r \sum_{n_1,n_2,m} \sum_\alpha$$
$$\times \left[\text{tr} \left((\tau_r \otimes s_0) \, Q_{n_1,n_1+m}^{\alpha\alpha}(\mathbf{x}) \right) \right]$$
$$\times \left[\text{tr} \left((\tau_r \otimes s_0) \, Q_{n_2+m,n_2}^{\alpha\alpha}(\mathbf{x}) \right) \right] \quad , \tag{2.24b}$$

$$\mathcal{A}_{int}^{(t)} = \frac{T\Gamma^{(t)}}{2} \int d\mathbf{x} \sum_{r=0,3} (-1)^r \sum_{n_1,n_2,m} \sum_\alpha \sum_{i=1}^3$$
$$\times \left[\text{tr} \left((\tau_r \otimes s_i) \, Q_{n_1,n_1+m}^{\alpha\alpha}(\mathbf{x}) \right) \right]$$
$$\times \left[\text{tr} \left((\tau_r \otimes s_i) \, Q_{n_2+m,n_2}^{\alpha\alpha}(\mathbf{x}) \right) \right] \quad , \tag{2.24c}$$

$$\mathcal{A}_{int}^{(c)} = \frac{T\Gamma^{(c)}}{2} \int d\mathbf{x} \sum_{r=1,2} \sum_{n_1,n_2,m} \sum_\alpha$$
$$\times \left[\text{tr} \left((\tau_r \otimes s_0) \, Q_{n_1,-n_1+m}^{\alpha\alpha}(\mathbf{x}) \right) \right]$$
$$\times \left[\text{tr} \left((\tau_r \otimes s_0) \, Q_{-n_2,n_2+m}^{\alpha\alpha}(\mathbf{x}) \right) \right] \quad . \tag{2.24d}$$

2. Properties of the Q-matrices

We now derive some useful properties of the matrix field Q. Since B, Eqs. (2.19), is self-adjoint under the adjoint operation defined in Eq. (2.14b) and denoted by

a superscript '+', so is Q, $Q^+ = Q$. Notice that Q^+ is distinct from the hermitian conjugate of Q, which we write as $Q^\dagger \equiv (Q^T)^*$, with the star denoting complex conjugation. The definition of the former adjoint implies $Q^+ = C^T Q^T C$, with

$$^{ij}C_{nm}^{\alpha\beta} = \delta_{nm}\,\delta_{\alpha\beta}\,c_{ij} \quad , \qquad (2.25)$$

and c defined in Eq. (2.14c). We thus have

$$Q = C^T\,Q^T\,C \quad . \qquad (2.26a)$$

In addition, a direct calculation of the hermitian conjugate of B, $B^\dagger \equiv \overline{B^T}$, reveals another constraint that is inherited by the Q, namely

$$Q^\dagger = -\Gamma\,Q\,\Gamma^{-1} \quad , \qquad (2.26b)$$

where the similarity transformation denoted by Γ has the property

$$(\Gamma\,Q\,\Gamma^{-1})_{nm} = Q_{-n-1,-m-1} \quad . \qquad (2.26c)$$

We now expand our matrix fields in the spin–quaternion basis defined after Eqs. (2.14),

$$Q_{12}(\mathbf{x}) = \sum_{r,i=0}^{3} (\tau_r \otimes s_i)\,{}_r^i Q_{12}(\mathbf{x}) \quad , \qquad (2.27a)$$

$$\tilde{\Lambda}_{12}(\mathbf{x}) = \sum_{r,i=0}^{3} (\tau_r \otimes s_i)\,{}_r^i \tilde{\Lambda}_{12}(\mathbf{x}) \quad . \qquad (2.27b)$$

where again $1 \equiv (n_1, \alpha_1)$, etc. In this basis, the relations expressed by Eq. (2.26a) imply the following symmetry properties,

$${}_r^0 Q_{12} = (-)^r\,{}_r^0 Q_{21} \quad , \quad (r = 0, 3) \quad , \qquad (2.28a)$$

$${}_r^i Q_{12} = (-)^{r+1}\,{}_r^i Q_{21} \quad , \quad (r = 0, 3;\; i = 1, 2, 3) \quad , \qquad (2.28b)$$

$${}_r^0 Q_{12} = {}_r^0 Q_{21} \quad , \quad (r = 1, 2) \quad , \qquad (2.28c)$$

$${}_r^i Q_{12} = -{}_r^i Q_{21} \quad , \quad (r = 1, 2;\; i = 1, 2, 3) \quad . \qquad (2.28d)$$

Together with the behavior under hermitian conjugation, Eq. (2.26b), this further implies

$${}_r^i Q_{12}^* = -{}_r^i Q_{-n_1-1,-n_2-1}^{\alpha_1\alpha_2} \quad , \quad (r = 0, 3) \quad , \qquad (2.29a)$$

$${}_r^i Q_{12}^* = {}_r^i Q_{-n_1-1,-n_2-1}^{\alpha_1\alpha_2} \quad , \quad (r = 1, 2) \quad , \qquad (2.29b)$$

Analogous relations hold for $\tilde{\Lambda}$ by virtue of the linear coupling between Q and $\tilde{\Lambda}$.

This completes the derivation of the action in terms of classical matrix fields. We reiterate that for disordered fermions, this action is adequate as long as one is interested in physics dominated by soft modes, while for clean systems, a complete theory would have to keep an infinite set of interaction constants. We also recall that, by means of the Q-matrices, we keep explicitly only density modes in both the particle–hole and the particle–particle

channels. While this is again physically well motivated in the disordered case, it is more problematic in the clean case. Due to the larger number of soft modes in the latter, our procedure means that in certain clean system soft modes have been integrated out. While this integrating–out procedure is exact, it leads to undesirable properties of the effective field theory as we will discuss below. In the next subsection we will find further indications for it being advantageous to derive somewhat different effective theories for clean and disordered fermions, respectively.

C. Saddle–point solutions and Gaussian approximation

The physical degrees of freedom are now given by the matrix elements of the Q-matrices, and the physical correlation functions for number and spin density fluctuations can be expressed in terms of Q-correlation functions. This implies that ultimately we will want to integrate out the ghost field $\tilde{\Lambda}$ in Eq. (2.21). Before doing so, however, we examine a saddle–point solution to the field theory defined by Eqs. (2.21) - (2.24), and the Gaussian fluctuations about this saddle point.

1. Fermi–liquid saddle–point solution

It is possible to develop a systematic theory for the expectation values of Q and $\tilde{\Lambda}$, rather than simply a saddle–point approximation for these quantities. We will come back to this point in Sec. III B 1 below. We further note that there are many saddle–point solutions, and that in general they have different symmetry properties. Physically these different solutions correspond to states with different broken symmetries such as Fermi–liquid states, magnetically ordered states, or superconducting states. For now we concentrate on a Fermi–liquid state. In Sec. III B 4 below, and elsewhere,[27] we will consider more exotic states.

The saddle–point condition is

$$\left.\frac{\delta\mathcal{A}}{\delta Q}\right|_{Q_{sp},\tilde{\Lambda}_{sp}} = \left.\frac{\delta\mathcal{A}}{\delta\tilde{\Lambda}}\right|_{Q_{sp},\tilde{\Lambda}_{sp}} = 0 \quad . \qquad (2.30)$$

The Fermi–liquid saddle–point solution is spatially uniform, diagonal in frequency and replica space, and there are no nonzero expectation values that would describe a spontaneous magnetization, or a BCS-like order parameter. In Fermi liquid–like phases, the saddle point values of both Q and $\tilde{\Lambda}$ have the structures,

$$\left.{}_r^i Q_{12}(\mathbf{x})\right|_{sp} = \delta_{12}\,\delta_{r0}\,\delta_{i0}\,Q_n \quad , \qquad (2.31a)$$

$$\left.{}_r^i \tilde{\Lambda}_{12}(\mathbf{x})\right|_{sp} = \delta_{12}\,\delta_{r0}\,\delta_{i0}\,\Lambda_n \quad , \qquad (2.31b)$$

and the average values of Q and $\tilde{\Lambda}$ have the same properties. Using Eqs. (2.21) - (2.31), we obtain the saddle point equations

$$Q_n = \frac{i}{2V} \sum_{\mathbf{p}} \left[i\omega_n - \mathbf{p}^2/2m + \mu - i\Lambda_n \right]^{-1} \quad , \quad (2.32a)$$

$$\Lambda_n = \frac{-2}{\pi N_F \tau_{rel}} Q_n - 4T\Gamma^{(s)} \sum_{m} e^{i\omega_m 0} Q_m \quad . \quad (2.32b)$$

Here $e^{i\omega_m 0}$ is the usual convergence factor that resolves the ambiguity presented by the equal-time Green function.[2] Notice that this saddle–point solution obeys all of the symmetry properties of the Q that were derived in Sec. II B 2. For a physical interpretation of Eqs. (2.32), we define a self-energy, $\Sigma_n = i\Lambda_n$. In terms of Σ, the Eqs. (2.32) can be rewritten,

$$\Sigma_n = \frac{1}{\pi N_F \tau_{rel}} \frac{1}{V} \sum_{\mathbf{p}} \left[i\omega_n - \mathbf{p}^2/2m + \mu - \Sigma_n \right]^{-1}$$

$$+ 2\Gamma^{(s)} T \sum_{m} e^{i\omega_m 0} \frac{1}{V} \sum_{\mathbf{p}} \left[i\omega_m - \mathbf{p}^2/2m \right.$$

$$\left. + \mu - \Sigma_m \right]^{-1} \quad . \quad (2.33)$$

This integral equation for the frequency dependent self–energy has a familiar structure. For vanishing interaction, $\Gamma^{(s)} = 0$, Eq. (2.33) reduces to the self–consistent Born approximation for the self–energy, or the inverse mean–free time, in a disordered fermion system. At zero disorder, on the other hand, Eq. (2.33) represents the Hartree–Fock approximation for the self–energy. The full Eq. (2.33) we will refer to as the disordered Hartree–Fock approximation for the self–energy. It interpolates smoothly between the clean and dirty cases.

In this context, a few remarks on ultraviolet divergencies and cutoffs are in order. In general dimensions, the sums or integrals in Eqs. (2.32) and (2.33) do not exist due to ultraviolet divergencies. Clearly, this is due to the long–wavelength and low–frequency approximations we have made by, for example, projecting onto density modes only. Any ultraviolet problems encountered in the theory are therefore artifacts, and should be cut off at some momentum scale k_0. In principle there are a number of 'large' momentum scales that are candidates for k_0, e.g. the Fermi wave number k_F, the Thomas–Fermi wave number k_{TF}, and the inverse elastic mean–free path $1/\ell$. For metallic densities, $k_F \approx k_{TF}$, while $k_F\ell >> 1$ for weakly disordered systems, and $k_F\ell \approx 1$ for strongly disordered ones. A popular choice is to use $k_0 = 2\pi/\ell$ in any disordered system, regardless of the strength of the disorder.[28] However, for weakly disordered systems this choice leads to qualitatively wrong answers for various nonuniversal results, while using $k_0 = k_F$ is qualitatively correct.[29] For the universal phenomena that will be considered in the present paper, on the other hand, the precise value of k_0 is not important, and any of the above choices are acceptable.

2. Gaussian approximation

We now examine the Gaussian fluctuations about the saddle point discussed above. To this end, we write Q and $\tilde{\Lambda}$ in Eqs. (2.21) - (2.24) as,

$$Q = Q_{sp} + \delta Q \quad , \quad (2.34a)$$

$$\tilde{\Lambda} = \tilde{\Lambda}_{sp} + \delta\tilde{\Lambda} \quad , \quad (2.34b)$$

and then expand to second or Gaussian order in the fluctuations δQ and $\delta\tilde{\Lambda}$. Denoting the constant saddle point contribution to the action by $\mathcal{A}_{sp}$, and the Gaussian action by $\mathcal{A}_G$, we have

$$\mathcal{A}[Q, \tilde{\Lambda}] = \mathcal{A}_{sp} + \mathcal{A}_G[Q, \tilde{\Lambda}] + \Delta\mathcal{A} \quad , \quad (2.35)$$

where $\Delta\mathcal{A}$ contains all terms of higher than quadratic order in the small quantities δQ and $\delta\tilde{\Lambda}$. The Gaussian action reads,

$$\mathcal{A}_G[Q, \tilde{\Lambda}] = \mathcal{A}_{dis}[\delta Q] + \mathcal{A}_{int}[\delta Q] + \frac{1}{4} \text{Tr} \left(G_{sp}\delta\tilde{\Lambda} G_{sp}\, \delta\tilde{\Lambda} \right)$$

$$+ \int d\mathbf{x}\, \text{tr} \left(\delta\tilde{\Lambda}(\mathbf{x})\, \delta Q(\mathbf{x}) \right) \quad , \quad (2.36a)$$

with G_{sp} the single–particle Green function in saddle–point approximation,

$$G_{sp}(\mathbf{p}, \omega_n) = \left[i\omega_n - \mathbf{p}^2/2m + \mu - \Sigma_n \right]^{-1} \quad , \quad (2.36b)$$

where Σ_n is given by the solution of Eq. (2.33). The spectrum of this saddle–point Green function determines the quantity N_F, see Eq. (2.46) below.

At this point it is important to notice that in the full action, Eq. (2.35), the symmetry properties expressed by Eqs. (2.28) are enforced by means of delta-function constraints. Truncating the theory at the Gaussian level spoils this property. Consequently, by using the constraints in different ways *before* the truncation one obtains different Gaussian theories. Here we opt to use the constraints to rewrite the theory entirely in terms of matrix elements δQ_{12} with $n_1 \geq n_2$. Correlation functions of matrix elements that do not obey these frequency restrictions are related to Eqs. (2.37) below by means of Eqs. (2.28). As we will see, this choice results in a Gaussian theory that reproduces results that are well-known from many-body perturbation theory. With this procedure, Eq. (2.36a) is a quadratic form that can be diagonalized using the techniques discussed in Ref. 16. One obtains

$$\left\langle {}^i_r(\delta Q)_{12}(\mathbf{p}_1)\, {}^j_s(\delta Q)_{34}(\mathbf{p}_2) \right\rangle_G = \frac{V}{16}\, \delta_{\mathbf{p}_1, -\mathbf{p}_2}\, \delta_{rs}\, \delta_{ij}$$

$$\times {}^i_r M^{-1}_{12,34}(\mathbf{p}_1) \,, \quad (2.37a)$$

$$\left\langle {}^i_r(\delta\bar{\Lambda})_{12}(\mathbf{p}_1)\, {}^j_s(\delta\bar{\Lambda})_{34}(\mathbf{p}_2) \right\rangle_G = \frac{V}{16}\, \delta_{\mathbf{p}_1, -\mathbf{p}_2}\, \delta_{rs}\, \delta_{ij}$$

$$\times {}^i_r N^{-1}_{12,34}(p_1) \,. \quad (2.37b)$$

9

Here $\langle\ldots\rangle_G$ denotes an average with a weight $\mathcal{A}_G$. In Eq. (2.37b), we have defined a correlation function for the fluctuations of a field, $\bar{\Lambda}$, that is decoupled from the Q-correlation function. $\bar{\Lambda}$ is defined by

$$\bar{\Lambda}_{12} = \frac{1}{2}\,\varphi_{n_1 n_2}\,\tilde{\Lambda}_{12} + Q_{12} \quad , \qquad (2.38a)$$

with

$$\varphi_{nm}(\mathbf{k}) = \frac{1}{V}\sum_{\mathbf{p}} G_{sp}(\mathbf{p},\omega_n)\,G_{sp}(\mathbf{p}+\mathbf{k},\omega_m) \quad . \quad (2.38b)$$

The matrix inverse M^{-1} for the particle-hole channel is given by,

$$\begin{aligned}
{}^{i}_{0,3}M^{-1}_{12,34}(\mathbf{p}) &= \delta_{1-2,3-4}\,{}^{i}_{0,3}I_{12}\Big[\delta_{13}\,\mathcal{D}_{n_1 n_2}(\mathbf{p})\\
&\quad -\delta_{\alpha_1\alpha_2}\,\delta_{\alpha_3\alpha_4}\,2T\Gamma^{(i)}\,\mathcal{D}^{(i)}_{n_1 n_2}(\mathbf{p})\,\mathcal{D}_{n_3 n_4}(\mathbf{p})\Big]\\
&\quad -\frac{4\delta_{r0}\delta_{i0}}{\pi N_F \tau_1}\,\delta_{12}\delta_{34}\mathcal{D}^{(s)}_{n_1 n_1}(\mathbf{p})\mathcal{D}^{(s)}_{n_3 n_3}(\mathbf{p}) \quad ,
\end{aligned}$$

$$(2.39a)$$

with

$$\begin{aligned}
{}^{i}_{0,3}I_{12} = 1 + \delta_{n_1 n_2}\big(1-2\delta_{r3}\big)\Big(\delta_{i0}-\sum_{j=1}^{3}\delta_{ij}\Big) \quad , \quad (2.39b)
\end{aligned}$$

and $\Gamma^{(0)} = -\Gamma^{(s)}$, and $\Gamma^{(1,2,3)} = \Gamma^{(t)}$. For the particle-particle channel, one finds,

$$\begin{aligned}
{}^{i}_{1,2}M^{-1}_{12,34}(\mathbf{p}) &= -\delta_{1+2,3+4}\Big[\delta_{13}\,{}^{i}_{1,2}I_{12}\,\mathcal{D}_{n_1 n_2}(\mathbf{p})\\
&\quad -\delta_{i0}\,\delta_{\alpha_1\alpha_2}\,\delta_{\alpha_3\alpha_4}\,4T\Gamma^{(c)}\mathcal{D}^{(c)}_{n_1 n_2}(\mathbf{p})\,\mathcal{D}_{n_3 n_4}(\mathbf{p})\Big] ,
\end{aligned}$$

$$(2.39c)$$

where

$$\begin{aligned}
{}^{i}_{1,2}I_{12} = 1 + \delta_{n_1 n_2}\Big(\delta_{i0}-\sum_{j=1}^{3}\delta_{ij}\Big) \quad . \quad (2.39d)
\end{aligned}$$

Similarly, one finds for the matrix inverse N^{-1} in the particle-hole channel,

$$\begin{aligned}
{}^{i}_{0,3}N^{-1}_{12,34}(\mathbf{p}) = -\delta_{13}\,\delta_{24}\,{}^{i}_{0,3}I_{12}\,\varphi_{n_1 n_2}(\mathbf{p}) \quad , \quad (2.39e)
\end{aligned}$$

and in the particle-particle channel,

$$\begin{aligned}
{}^{i}_{1,2}N^{-1}_{12,34}(\mathbf{p}) = \delta_{13}\,\delta_{24}\,{}^{i}_{1,2}I_{12}\,\varphi_{n_1 n_2}(\mathbf{p}) \quad , \quad (2.39f)
\end{aligned}$$

In the preceding equations,

$$\mathcal{D}_{n_1 n_3}(\mathbf{p}) = \varphi_{n_1 n_2}(\mathbf{p})\left[1-\frac{1}{\pi N_F \tau_{rel}}\,\varphi_{n_1 n_2}(\mathbf{p})\right]^{-1} \quad ,$$

$$(2.40a)$$

and,

$$\mathcal{D}^{(i)}_{n_1 n_2}(\mathbf{p}) = \mathcal{D}_{n_1 n_2}(\mathbf{p})\left[1+2T\Gamma^{(i)}\right.$$
$$\left.\times\sum_{n}\mathcal{D}_{n,n+n_2-n_1}(\mathbf{p})\right]^{-1} \quad , \quad (2.40b)$$

with $\mathcal{D}^{(0)} = \mathcal{D}^{(s)}$, and $\mathcal{D}^{(1)} = \mathcal{D}^{(2)} = \mathcal{D}^{(3)} = \mathcal{D}^{(t)}$, and

$$\mathcal{D}^{(c)}_{n_1 n_2}(\mathbf{p}) = \mathcal{D}_{n_1 n_2}(\mathbf{p})\left[1+2T\Gamma^{(c)}\right.$$
$$\left.\times\sum_{n}\mathcal{D}_{n,n_1+n_2-n}(\mathbf{p})\right]^{-1} \quad .$$

$$(2.40c)$$

D. Physical correlation functions

We conclude this section by interpreting the results obtained in the preceding subsection, and expressing some physical quantities of interest in terms of Q-correlations.

First of all, $\varphi_{nm}(\mathbf{k})$, Eq. (2.38b), is related to a disordered Lindhard function. We write

$$\varphi_{nm}(\mathbf{k}) = \Theta(nm)\Phi + \Theta(-nm)\varphi(\mathbf{k},\Omega_{n-m}) \quad , \quad (2.41)$$

with $\Omega_n = 2\pi T n$ a bosonic Matsubara frequency. For both clean and disordered systems, Φ approaches a finite constant in the limit of vanishing wave number and frequency. On the other hand, the behavior of φ in the same limit depends qualitatively on the presence or absence of disorder. For dirty systems, to leading order as $\tau_{rel}\to\infty$,

$$\varphi(|\mathbf{k}|\to 0, \Omega_n\to 0) = \pi N_F \tau_{rel} \quad , \quad (2.42)$$

while for clean systems, φ is singular in the long–wavelength, low–frequency limit. For example, in $d=3$,

$$\varphi(|\mathbf{k}|\to 0, \Omega_n\to 0) = \frac{i\pi N_F}{2|\mathbf{k}|v_F}\,\mathrm{sgn}\,(\Omega_n)\ln\frac{i\Omega_n+|\mathbf{k}|v_F}{i\Omega_n-|\mathbf{k}|v_F} \quad .$$

$$(2.43)$$

Comparing Eqs. (2.42) and (2.43), we see that the $1/|\mathbf{k}|$ and $1/\Omega_n$ singularities that occur in the clean case are cut off by a mean–free path, and a relaxation time, respectively, in the disordered case. This implies that the $\bar{\Lambda}$-$\bar{\Lambda}$ correlation function is soft for clean systems, and massive for dirty ones. This will turn out to be a fundamental difference between clean and dirty systems, which reflects the fact that single particle excitations are soft in a clean system. These observations indicate that for the further development of the theory it will be useful to treat clean and disordered systems in somewhat different ways.

φ determines the propagator $\mathcal{D}$, Eq. (2.40a). For finite disorder, $n_1 n_2 \leq 0$, and in the long–wavelength, low–frequency limit one finds

$$\mathcal{D}_{n_1 n_2}(\mathbf{p}) = \frac{\pi N_F}{D\mathbf{p}^2 + |\Omega_{n_1 - n_2}|} \quad , \qquad (2.44)$$

with $D = v_F^2 \tau_{rel}/d$ a bare or Boltzmann diffusion coefficient.[30] The propagators given by Eqs. (2.37a) and (2.39a) for $n_1 n_2 < 0$ are then the soft or massless propagators of the nonlinear sigma–model that is used to describe the metal–insulator transition in disordered interacting electronic systems.[15,16]

Second, the physical correlation functions of the single–particle spectral density, the particle number density, and the spin density, can be easily expressed in terms of the Q-correlation functions. By keeping an appropriate source in the action while performing the transformation to Q-variables, we obtain an expression for the density of states, N, as a function of energy or frequency ω,

$$N(\epsilon_F + \omega) = \frac{4}{\pi} \text{Re} \left\langle {}_0^0 Q_{nn}(\mathbf{x}) \right\rangle \Big|_{i\omega_n \to \omega + i0} \quad . \qquad (2.45)$$

In saddle point approximation, we have for the density of states at the Fermi level,

$$N_F = \frac{-2}{\pi} \frac{1}{V} \sum_{\mathbf{p}} \text{Im} \, G_{sp}(\mathbf{p}, i\omega_n \to i0) \quad . \qquad (2.46)$$

This defines the density of states that has been used throughout as an energy normalization factor. Similarly, one finds that the number density susceptibility, χ_n, and the spin density susceptibility, χ_s, are given by,

$$\chi^{(i)}(\mathbf{q}, \Omega_n) = 16T \sum_{m_1, m_2} \sum_{r=0,3} \left\langle {}_r^i(\Delta Q)_{m_1 - n, m_1}^{\alpha\alpha}(\mathbf{q}) \right.$$
$$\left. \times {}_r^i(\Delta Q)_{m_2, m_2 + n}^{\alpha\alpha}(-\mathbf{q}) \right\rangle \quad , \quad (2.47)$$

with $\Delta Q = Q - \langle Q \rangle$, and $\chi^{(0)} = \chi_n$ and $\chi^{(1,2,3)} = \chi_s$. By substituting the Gaussian propagator we see that the Gaussian theory yields χ_n and χ_s in a disordered RPA-like approximation. Notice that, for weak disorder, the last term on the right–hand side of Eq. (2.39a) is smaller than the first two terms by at least a factor of $1/\tau_1$, and therefore does not appear in the usual resummation schemes of many–body perturbation theory. Neglecting it, one finds, for small $|\mathbf{q}|$ and Ω_n,

$$\chi^{(i)}(\mathbf{q}, \Omega_n) = \chi_{st}^{(i)} \frac{-D^{(i)}\mathbf{q}^2}{|\Omega_n| + D^{(i)}\mathbf{q}^2} \quad , \qquad (2.48a)$$

where

$$\chi_{st}^{(i)} = N_F / \left(1 - N_F \Gamma^{(i)}\right) \quad , \qquad (2.48b)$$

is the static susceptibility in the long–wavelength limit, renormalized by the interaction, and

$$D^{(i)} = D \left(1 - N_F \Gamma^{(i)}\right) \quad , \qquad (2.48c)$$

is the renormalized diffusion coefficient. Here we have used the identity $T \sum_n \mathcal{D}_{nn}(\mathbf{q} \to 0) = -N_F/2$, which

expresses particle number conservation. Notice that for $i = 1, 2, 3$, Eqs. (2.48) describe a disordered Stoner criterion: For $N_F \Gamma^{(t)} = 1$, the spin diffusion coefficient vanishes, and the static spin susceptibility diverges. In the clean limit one recovers RPA proper,

$$\chi^{(i)}(\mathbf{q}, \Omega_n) = 2\chi_0(\mathbf{q}, \Omega_n) \left[1 + \Gamma^{(i)} 2\chi_0(\mathbf{q}, \Omega_n)\right]^{-1} \quad ,$$
$$(2.49)$$

where $\chi_0(\mathbf{q}, \Omega_n) = T \sum_m \varphi_{m-n,m}(\mathbf{q})$ is the free electron density susceptibility per spin.

III. THE DISORDERED FERMION SYSTEM

In this section we concentrate on the disordered system, and in particular perform a symmetry analysis in order to identify the soft modes in our system. This will allow us to explicitly separate the soft modes from the massive ones, and to formulate an effective field theory for which the soft modes remain manifestly soft to all orders in perturbation theory. In Appendix B we point out that all of the considerations and results of this section have precise analogies within the ϕ^4–representation of the $O(N)$ Heisenberg model. These analogies are often helpful, since ϕ^4–theory is technically much simpler than the matrix field theory under consideration here, and hence it displays the basic structure of the symmetry analysis and its consequences more clearly.

A. Symmetry analysis

1. Basic transformation properties

Let us perform a symmetry analysis of our field theory. To this end, we start with the action in fermionic form, formulated in terms of bispinors, Eqs. (2.16) - (2.18). Our plan of attack is to consider the various terms in the action separately. The crucial symmetries will involve the free electron part, $\tilde{S}_0$, and the disorder part, $\tilde{S}_{dis}$. The interaction part, $\tilde{S}_{int}$, Eqs. (2.17), will turn out to always be effectively proportional to a frequency, at least near the disordered Fermi–liquid FP that we will identify below. Since we are interested in low–frequency effects, this will ultimately imply that $\tilde{S}_{int}$ does not change the conclusions with respect to which two–particle excitations are soft or massless that are reached by considering the noninteracting action.

For zero external frequency, $i\omega_n = 0$ in Eq. (2.16a), the free fermion action is invariant under transformations that leave invariant the expression

$$\sum_n \sum_\alpha \text{tr} \left((\eta^+)_n^\alpha \otimes \eta_n^\alpha\right) \equiv (\eta, \eta) \quad , \qquad (3.1)$$

with (η, η) a generalization of the scalar product defined in Eq. (2.15) that includes summation over the frequency and replica indices. From the structure of the disorder part of the action, Eq. (2.16b), it is clear that $\tilde{S}_{dis}$ is also invariant under transformations that leave invariant (η, η).

Now let η be transformed by means of an operator $\hat{T}$: $\eta \to \hat{T}\eta$. Remembering that η^+ is related to η by means of the charge conjugation matrix C, Eq. (2.25), and using $C^T = C^{-1}$, we find that in order to leave (η, η) invariant, $\hat{T}$ must obey

$$\hat{T}^T C \hat{T} = C \quad . \tag{3.2}$$

For a system with $2N$ frequency labels (N positive ones, including 0, and N negative ones), and n replicas, Eq. (3.2) defines a representation of the symplectic group $\mathrm{Sp}(8Nn, \mathcal{C})$ over the complex numbers $\mathcal{C}$.[31] Throughout this section we will denote the real, complex, and quaternion number fields by $\mathcal{R}$, $\mathcal{C}$, and $\mathcal{Q}$, respectively. In what follows we will need to consider only small subgroups of the $8Nn(8Nn + 1)$-parameter group $\mathrm{Sp}(8Nn, \mathcal{C})$.

Now we ask what transformations of our composite variables Q correspond to the above transformations of the bispinors. Since the Q-matrix is constrained to the bilinear fermion field $B = \eta^+ \otimes \eta$, see Eq. (2.20), $\eta \to \hat{T}\eta$ implies $Q \to TQT^{-1}$, with $T = C\hat{T}C^T$. The ghost field $\tilde{\Lambda}$ couples linearly to Q and B, and therefore transforms like Q, i.e. by means of the matrices T. Together with the $\hat{T}$, the transformations T also obey Eq. (3.2), and therefore also form a representation of the group $\mathrm{Sp}(8Nn, \mathcal{C})$.

Now we consider a specific transformation, namely a rotation in frequency space given by,

$$
\begin{aligned}
{}_r^i \hat{T}_{nm}^{\alpha\beta} = \delta_{i0}\, \delta_{r0}\, \delta_{\alpha\beta} &\left\{ \delta_{nm} \left[1 + (\delta_{nn_1} + \delta_{nn_2})(\cos\theta - 1) \right] \right. \\
&\left. + (\delta_{nn_1}\delta_{mn_2} - \delta_{nn_2}\delta_{mn_1}) \sin\theta \right\} \\
\equiv\; &\delta_{i0}\, \delta_{r0}\, \delta_{\alpha\beta}\, \hat{t}_{nm} \quad .
\end{aligned} \tag{3.3}
$$

This transforms a pair of spinors with frequency labels n_1 and n_2, respectively, into linear combinations of the same pair, with a mixing angle θ. These transformations obey Eq. (3.2), and thus are elements of $\mathrm{Sp}(8Nn, \mathcal{C})$. For fixed n_1 and n_2 they represent an SO(2) subgroup of $\mathrm{Sp}(8Nn, \mathcal{C})$. The corresponding transformations T of the Q and $\tilde{\Lambda}$-matrices are identical with the $\hat{T}$, $T = \hat{T}$. Under an infinitesimal element of this subgroup, the Q-matrices transform like

$$Q_{nm} \to Q_{nm} + \delta Q_{nm} \quad , \tag{3.4a}$$

with

$$
\begin{aligned}
\delta Q_{nm} = \big(\delta_{nn_1} Q_{n_2 m} - \delta_{nn_2} Q_{n_1 m} + \delta_{mn_1} Q_{nn_2} \\
- \delta_{mn_2} Q_{nn_1} \big)\, \theta + O(\theta^2) \quad . \tag{3.4b}
\end{aligned}
$$

Here we have shown only the frequency indices, since all other degrees of freedom are unaffected by the transformation. The $\tilde{\Lambda}$-matrices transform accordingly.

Of the various pieces of the action in Q-$\tilde{\Lambda}$ formulation, Eq. (2.21), $\mathcal{A}_{dis}$ and $\int d\mathbf{x}\, \mathrm{tr}\, (\tilde{\Lambda} Q)$ are invariant under the above transformation, but $\mathrm{Tr}\, \ln(G_0^{-1} - i\tilde{\Lambda})$ and $\mathcal{A}_{int}$ are not. For now, let us consider the transformation of the former:

$$\mathrm{Tr}\, \ln\left(G_0^{-1} - i\tilde{\Lambda} \right) \to \mathrm{Tr}\, \ln\left(G_0^{-1} - i\tilde{\Lambda} \right) + \theta\, \mathrm{Tr}\, (G\, \delta i\omega) \quad , \tag{3.5a}$$

with $G \equiv (G_0^{-1} - i\tilde{\Lambda})^{-1}$, and

$$
{}_r^i (\delta i\omega)_{nm}^{\alpha\beta} = \delta_{i0}\, \delta_{r0}\, \delta_{\alpha\beta}\, (\delta_{nn_1}\delta_{mn_2} + \delta_{nn_2}\delta_{mn_1})\, i\Omega_{n_1 - n_2} \,. \tag{3.5b}
$$

The action of the noninteracting Fermi system therefore transforms like

$$\mathcal{A}_0 = \mathcal{A} - \mathcal{A}_{int} \to \mathcal{A}_0 + \delta\mathcal{A}_0 = \mathcal{A}_0 + \frac{\theta}{2}\, \mathrm{Tr}\, (G\, \delta i\omega) \quad . \tag{3.6}$$

We will consider the transformation properties of $\mathcal{A}_{int}$ in Sec. III A 3 below.

2. A Ward identity for noninteracting electrons

Because of the complexity of the interacting case, and for the reasons noted at the beginning of the previous subsection, we first discuss the noninteracting system. Let us introduce a source J for the Q-fields, and consider the partition function

$$Z[J] = \int D[Q]\, D[\tilde{\Lambda}]\; e^{\mathcal{A}_0 + \int d\mathbf{x}\, \mathrm{tr}\, \left(J(\mathbf{x})\, Q(\mathbf{x}) \right)} \quad . \tag{3.7}$$

By performing an infinitesimal transformation, Eq. (3.4b), on the Q and $\tilde{\Lambda}$-fields, one obtains from Eq. (3.7) the identity,

$$
\begin{aligned}
0 = \int D[Q]\, D[\tilde{\Lambda}]\; &\left[\delta\mathcal{A}_0 + \int d\mathbf{x}\, \mathrm{tr}\, \left(J(\mathbf{x})\, \delta Q(\mathbf{x}) \right) \right] \\
&\times e^{\mathcal{A}_0 + \int d\mathbf{x}\, \mathrm{tr}\, \left(J(\mathbf{x})\, Q(\mathbf{x}) \right)} \quad . \tag{3.8a}
\end{aligned}
$$

Differentiating this identity with respect to ${}_0^0 J_{n_2 n_1}^{\beta\alpha}$ and putting $J = 0$ yields

$$0 = \left\langle \delta\mathcal{A}_0\; {}_0^0 Q_{n_1 n_2}^{\alpha\beta}(\mathbf{x}) \right\rangle + \left\langle {}_0^0 (\delta Q)_{n_1 n_2}^{\alpha\beta}(\mathbf{x}) \right\rangle \quad , \tag{3.8b}$$

where $\langle \ldots \rangle$ denotes an average by means of the action $\mathcal{A}_0$. From the explicit expression for $\delta\mathcal{A}_0$, Eq. (3.6), we see that the first average is essentially a frequency times $\langle G\, Q \rangle$. By using the identity[32] $\langle G\, Q \rangle = -2i\langle Q^2 \rangle$ we finally obtain

$$8\Omega_{n_1-n_2} \sum_\gamma \int d\mathbf{y} \, \left\langle {}_0^0 Q_{n_1 n_2}^{\gamma\gamma}(\mathbf{y}) \, {}_0^0 Q_{n_1 n_2}^{\alpha\beta}(\mathbf{x}) \right\rangle$$

$$= \left\langle {}_0^0 Q_{n_1 n_1}^{\alpha\beta}(\mathbf{x}) \right\rangle - \left\langle {}_0^0 Q_{n_2 n_2}^{\alpha\beta}(\mathbf{x}) \right\rangle \quad . \qquad (3.9)$$

This is the desired Ward identity for noninteracting electrons,[33,17,18] which relates a two–point Q-correlation function to the one–point function. For $\mathrm{sgn}\, n_1 = \mathrm{sgn}\, n_2$, the right–hand side vanishes as $n_1 - n_2 \to 0$, and the identity tells us nothing interesting. However, for $\mathrm{sgn}\, n_1 \neq \mathrm{sgn}\, n_2$ we see from Eq. (2.45) that the right–hand side approaches $\pi N(\epsilon_F)/2$, which is nonzero everywhere inside the band. The correlation function on the left–hand side must therefore diverge. Due to rotational invariance in replica space we have $\langle Q^{\alpha\beta} \rangle \sim \delta_{\alpha\beta}$, and $\langle Q^{\alpha\beta} Q^{\gamma\delta} \rangle \sim \delta_{\alpha\gamma}\delta_{\beta\delta}$. We finally obtain, for $n_1 n_2 < 0$ and $|\Omega_{n_1-n_2}| \to 0$,

$$\left\langle {}_0^0 Q_{n_1 n_2}^{\alpha\alpha}(\mathbf{k}) \, {}_0^0 Q_{n_1 n_2}^{\alpha\alpha}(-\mathbf{k}) \right\rangle \Big|_{\mathbf{k}=0} = \frac{\pi N(\epsilon_F)}{16|\Omega_{n_1-n_2}|} \quad . \quad (3.10)$$

Here $N(\epsilon_F)$ is the exact density of states at the Fermi level of the noninteracting electron system. The salient point is that, as long as $N(\epsilon_F) > 0$, the Q-Q correlation function at zero momentum diverges like $1/|\Omega_{n_1-n_2}|$. Equations (2.37a), (2.39a), and (2.44) show that the Gaussian propagator has this property. The Ward identity ensures that it holds to all orders in perturbation theory. We have therefore identified ${}_0^0 Q_{n_1 n_2}^{\alpha\alpha}$ for $n_1 n_2 < 0$ as a soft mode. Rotational symmetry in replica space implies that the ${}_0^0 Q_{nm}^{\alpha\beta}$, $nm < 0$, $\alpha \neq \beta$, are also soft. This can also be seen by applying the above rotation procedure in frequency space directly to spinor pairs with different replica indices.

From the Gaussian propagator, Eqs. (2.37a), (2.39), as well as from physical considerations, one expects all channels to be soft, not just the $r = 0, i = 0$ channel. This is indeed the case. These additional soft modes are not controlled by separate Ward identities, but rather are related to Eq. (3.9) by additional symmetries of the action. In order to see this, we consider transformations

$$T_{nm}^{\alpha\beta} = \delta_{\alpha\beta} \, \delta_{nm} \left[\delta_{nn_2} x_r \, (\tau_r \otimes s_i) + (1 - \delta_{nn_2})(\tau_0 \otimes s_0) \right] ,$$
$$(3.11a)$$

with $x_0 = x_3 = 1$, and $x_1 = x_2 = i$, and n_2 some fixed frequency index. These T obey Eq. (3.2), and in addition they are unitary. It is easy to check that under these transformations Q transforms like

$$ {}_0^0 Q_{nn_2} \to {}_r^i Q_{nn_2} \quad (n \neq n_2) \quad . \qquad (3.11b)$$

If the action is invariant under such a transformation for given i, r, then $\langle {}_r^i Q \, {}_r^i Q \rangle = \langle {}_0^0 Q \, {}_0^0 Q \rangle$, and hence the two–point Q-correlations in the i, r channel are also soft. Obviously, the noninteracting actions is indeed invariant under these transformations, so the two–point correlations are soft in all channels.

Note that Eqs. (3.11) imply that the spin–singlet ($i = 0$) and spin–triplet ($i = 1, 2, 3$) channels are equal, and so are the particle–hole ($r = 0, 3$) and particle–particle ($r = 1, 2$) channels. In particular, they provide a technical explanation for why the so–called Cooperons, i.e. the two–point Q-correlation functions in the particle–particle channel, are soft modes. We also mention that in the clean case, there still is only one basic Ward identity, but many more additional symmetries than in a disordered system.[27] These additional symmetries result in many additional soft modes. This feature leads to the somewhat paradoxical fact that it is easier to derive an effective field theory for disordered electron systems than for clean ones.

3. A Ward identity for interacting electrons

Now we add the electron–electron interaction to our considerations. For simplicity, we restrict ourselves to a discussion of the particle–hole spin–singlet interaction; the discussion of the remaining interaction channels proceeds analogously.

Let us consider again an infinitesimal element of the rotations in frequency space, Eq. (3.3). From Eqs. (2.24b) and (3.4b) we find that the particle–hole spin–singlet part of the action transforms like $\mathcal{A}_{int}^{(s)} \to \mathcal{A}_{int}^{(s)} + \delta\mathcal{A}_{int}^{(s)}$, with

$$\delta\mathcal{A}_{int}^{(s)} = 16\Gamma^{(s)} \int d\mathbf{x} \sum_\gamma \sum_{r=0,3} T \sum_{n_1' n_2'} {}_r^0 Q_{n_1' n_2'}^{\gamma\gamma}(\mathbf{x})$$

$$\times \left[{}_r^0 Q_{n_1+(n_1'-n_2'),n_2}^{\gamma\gamma}(\mathbf{x}) - {}_r^0 Q_{n_1,n_2-(n_1'-n_2')}^{\gamma\gamma}(\mathbf{x}) \right.$$

$$\left. - {}_r^0 Q_{n_2+(n_1'-n_2'),n_1}^{\gamma\gamma}(\mathbf{x}) + {}_r^0 Q_{n_2,n_1-(n_1'-n_2')}^{\gamma\gamma}(\mathbf{x}) \right] \theta \quad . \quad (3.12)$$

Now we follow the same steps that led to Eq. (3.8b), except that we differentiate with respect to ${}_0^0 J_{n_4 n_3}^{\beta\alpha}$ with $n_3 > 0, n_4 < 0$. This yields

$$- \left\langle {}_0^0 \left(\delta Q(\mathbf{x}) \right)_{n_3 n_4}^{\alpha\beta} \right\rangle = \left\langle \delta\mathcal{A}_0 \, {}_0^0 Q_{n_3 n_4}^{\alpha\beta}(\mathbf{x}) \right\rangle$$

$$+ \left\langle \delta\mathcal{A}_{int} \, {}_0^0 Q_{n_3 n_4}^{\alpha\beta}(\mathbf{x}) \right\rangle \quad . \quad (3.13)$$

Choosing $n_1 > 0, n_2 < 0$, and using $\langle Q_{nm} \rangle \sim \delta_{nm}$ as well as Eq. (3.12), we obtain the Ward identity in the form,

$$W_{int} + 8\Omega_{n_1-n_2} \sum_\gamma \int d\mathbf{y} \, \left\langle {}_0^0 Q_{n_1 n_2}^{\gamma\gamma}(\mathbf{y}) \, {}_0^0 Q_{n_3 n_4}^{\alpha\beta}(\mathbf{x}) \right\rangle$$

$$= \delta_{n_1 n_3} \delta_{n_2 n_4} \left(\left\langle {}_0^0 Q_{n_1 n_1}^{\alpha\beta}(\mathbf{x}) \right\rangle - \left\langle {}_0^0 Q_{n_2 n_2}^{\alpha\beta}(\mathbf{x}) \right\rangle \right) \quad ,$$
$$(3.14a)$$

where,

$$W_{int} = -16\Gamma^{(s)} \sum_\gamma \int d\mathbf{y} \sum_{r=0,3} T \sum_{n_1' n_2'} \left\langle {}_0^0 Q_{n_3 n_4}^{\alpha\beta}(\mathbf{x}) \right.$$

$$\times \, {}^{0}_{r}Q^{\gamma\gamma}_{n'_1 n'_2}(\mathbf{y}) \left[{}^{0}_{r}Q^{\gamma\gamma}_{n_1, n_2-(n'_1-n'_2)}(\mathbf{y}) \right.$$
$$- {}^{0}_{r}Q^{\gamma\gamma}_{n_1+(n'_1-n'_2), n_2}(\mathbf{y}) + {}^{0}_{r}Q^{\gamma\gamma}_{n_2+(n'_1-n'_2), n_1}(\mathbf{y})$$
$$\left. - {}^{0}_{r}Q^{\gamma\gamma}_{n_2, n_1-(n'_1-n'_2)}(\mathbf{y}) \right] \Big\rangle \quad ,$$
$$(3.14\text{b})$$

with $n_1, n_3 > 0$, and $n_2, n_4 < 0$.

We now need to analyze the 3-point correlation functions in Eq. (3.14b), and ascertain that the $Q_{n_1 n_2}$ with $n_1 n_2 < 0$ remain soft modes in their presence. To this end we first split up the Q into their averages, and fluctuations: $Q_{nm} = \langle Q_{nm} \rangle + (\Delta Q)_{nm} = \delta_{nm} \langle Q_{nn} \rangle + (\Delta Q)_{nm}$. Since $n_1 \neq n_2$, and $n_3 \neq n_4$, we have $\langle Q_{n_1 n_2} \rangle = \langle Q_{n_3 n_4} \rangle = 0$. Furthermore, if we put $n'_1 = n'_2$ in the 3-point functions, then the expression in square brackets in Eq. (3.14b) vanishes, so effectively we also have $\langle Q_{n'_1 n'_2} \rangle = 0$. Equation (3.14a) then takes the form,

$$W_{int} + 8\Omega_{n_1-n_2} \sum_{\gamma} \int d\mathbf{y} \, \Big\langle {}^{0}_{0}(\Delta Q)^{\gamma\gamma}_{n_1 n_2}(\mathbf{y})$$
$$\times {}^{0}_{0}(\Delta Q)^{\alpha\beta}_{n_3 n_4}(\mathbf{x}) \Big\rangle$$
$$= \delta_{n_1 n_3} \delta_{n_2 n_4} \left(\Big\langle {}^{0}_{0}Q^{\alpha\beta}_{n_1 n_1}(\mathbf{x}) \Big\rangle - \Big\langle {}^{0}_{0}Q^{\alpha\beta}_{n_2 n_2}(\mathbf{x}) \Big\rangle \right) \quad ,$$
$$(3.14\text{c})$$

and W_{int} can be written,

$$W_{int} = -16\Gamma^{(s)} \sum_{\gamma} \left(\Big\langle {}^{0}_{0}Q^{\gamma\gamma}_{n_1 n_1}(\mathbf{x}) \Big\rangle - \Big\langle {}^{0}_{0}Q^{\gamma\gamma}_{n_2 n_2}(\mathbf{x}) \Big\rangle \right)$$
$$\times T \sum_{n'_1 n'_2} \left(\delta_{n'_1-n'_2, n_2-n_1} + \delta_{n'_1-n'_2, n_1-n_2} \right)$$
$$\times \int d\mathbf{y} \, \Big\langle {}^{0}_{0}(\Delta Q)^{\gamma\gamma}_{n'_1 n'_2}(\mathbf{y}) \, {}^{0}_{0}(\Delta Q)^{\alpha\beta}_{n_3 n_4}(\mathbf{x}) \Big\rangle$$
$$-16\Gamma^{(s)} \sum_{\gamma} \int d\mathbf{y} \sum_{r=0,3} T \sum_{n'_1 n'_2}$$
$$\times \Big\langle {}^{0}_{0}(\Delta Q)^{\alpha\beta}_{n_3 n_4}(\mathbf{x}) \, {}^{0}_{r}(\Delta Q)^{\gamma\gamma}_{n'_1 n'_2}(\mathbf{y})$$
$$\times \left[{}^{0}_{r}(\Delta Q)^{\gamma\gamma}_{n_1, n_2-(n'_1-n'_2)}(\mathbf{y}) \right.$$
$$- {}^{0}_{r}(\Delta Q)^{\gamma\gamma}_{n_1+(n'_1-n'_2), n_2}(\mathbf{y})$$
$$+ {}^{0}_{r}(\Delta Q)^{\gamma\gamma}_{n_2+(n'_1-n'_2), n_1}(\mathbf{y})$$
$$\left. - {}^{0}_{r}(\Delta Q)^{\gamma\gamma}_{n_2, n_1-(n'_1-n'_2)}(\mathbf{y}) \right] \Big\rangle \quad ,$$
$$(3.14\text{d})$$

Next we remember that the composite variables Q correspond to products of fermions variables, e.g. $Q \sim \bar{\psi}\psi$ in the particle–hole channel. Specifically,

$$ {}^{0}_{0}Q_{n_1 n_2} \sim \frac{i}{2} \sum_{\sigma} \left(\bar{\psi}_{n_1,\sigma}\psi_{n_2,\sigma} + \bar{\psi}_{n_2,\sigma}\psi_{n_1,\sigma} \right) \quad , \quad (3.15\text{a})$$

$$ {}^{0}_{3}Q_{n_1 n_2} \sim \frac{-1}{2} \sum_{\sigma} \left(\bar{\psi}_{n_1,\sigma}\psi_{n_2,\sigma} - \bar{\psi}_{n_2,\sigma}\psi_{n_1,\sigma} \right) \quad , \quad (3.15\text{b})$$

Therefore, writing the correlation functions in terms of $\Delta Q \sim \bar{\psi}\psi - \langle \bar{\psi}\psi \rangle$ makes them Q-irreducible, but not ψ-irreducible. For instance, the 2-point function in Eqs. (3.14c) and (3.14d) has the structure,

$$\sum_{\gamma} \int d\mathbf{y} \, \Big\langle {}^{0}_{0}(\Delta Q)^{\gamma\gamma}_{n_1 n_2}(\mathbf{y}) \, {}^{0}_{0}(\Delta Q)^{\alpha\beta}_{n_3 n_4}(\mathbf{x}) \Big\rangle$$
$$= \delta_{\alpha\beta} \left[\delta_{n_1 n_3} \delta_{n_2 n_4} X^{(1)}_{n_1 n_2} + \delta_{n_1-n_2, n_3-n_4} X^{(2)}_{n_1 n_2, n_3 n_4} \right] \quad .$$
$$(3.16)$$

The replica structure of this identity is such that $X^{(1)}$ and $X^{(2)}$ are replica independent.[34] We can therefore suppress the replica index in what follows, as we have done already in Eqs. (3.15). The frequency structure of Eq. (3.16) is very general: By virtue of Eq. (3.15a), the full 2-point function is automatically proportional to $\delta_{n_1-n_2, n_3-n_4}$ due to time translational invariance, and the disconnected part is in addition proportional to $\delta_{n_1 n_3} \delta_{n_2 n_4}$. Using Eqs. (3.15) in the 3-point Q-correlations shows that they also contain parts that are proportional to $\delta_{n_1 n_3} \delta_{n_2 n_4}$, and others that are proportional to $\delta_{n_1-n_3, n_2-n_4}$. This suggests to break up the Ward identity into two separate pieces with these two frequency structures, respectively.

By collecting the above arguments, we find that the coefficients of $\delta_{n_1 n_3} \delta_{n_2 n_4}$ in the Ward identity obey the relation,

$$\Big\langle {}^{0}_{0}Q^{\alpha\beta}_{n_1 n_1}(\mathbf{x}) \Big\rangle - \Big\langle {}^{0}_{0}Q^{\alpha\beta}_{n_2 n_2}(\mathbf{x}) \Big\rangle = \delta_{\alpha\beta} \Big[8\Omega_{n_1-n_2} X^{(1)}_{n_1 n_2}$$
$$- N_F \left(N_F \Gamma^{(s)} \right)^2 Y^{(1)}_{n_1 n_2} \Big] \quad . \quad (3.17\text{a})$$

Here $Y^{(1)}$ is related to the piece of the 3-point correlation functions that is proportional to $\delta_{n_1 n_3} \delta_{n_2 n_4}$. An explicit calculation yields,

$$Y^{(1)}_{n_1 n_2} = \frac{-8}{N_F^3 \Gamma^{(s)}} \int d\mathbf{y} \, T \sum_{n'_1 n'_2}$$
$$\times \Big\{ G(\mathbf{x}, \mathbf{y}; \omega_{n_1}) \Big\langle \left(\bar{\psi}_{n'_1}(\mathbf{y}) s_0 \psi_{n'_2}(\mathbf{y}) \right)$$
$$\times \left((\bar{\psi}_{n_2-n'_1+n'_2}(\mathbf{y}) s_0 \psi_{n_2}(\mathbf{x})) \right) \Big\rangle^c$$
$$- G(\mathbf{x}, \mathbf{y}; \omega_{n_2}) \Big\langle \left(\bar{\psi}_{n'_1}(\mathbf{y}) s_0 \psi_{n'_2}(\mathbf{y}) \right)$$
$$\times \left((\bar{\psi}_{n_1-n'_1+n'_2}(\mathbf{y}) s_0 \psi_{n_1}(\mathbf{x})) \right) \Big\rangle^c \Big\}_{dis} \quad .$$
$$(3.17\text{b})$$

Here the $\langle \bar{\psi}\psi\bar{\psi}\psi \rangle^c$ are ψ-connected correlation functions, with the cumulant taken with respect to the quantum mechanical expectation value only, and $G(\mathbf{x}, \mathbf{y}; \omega_n) =$

$\langle \bar{\psi}_{n,\sigma}(\mathbf{x})\,\psi_{n,\sigma}(\mathbf{y})\rangle$ is the Green function for a given disorder configuration. For $Tn_1 \to +0$ and $Tn_2 \to -0$, so that $\Omega_{n_1-n_2} \to 0$, the left–hand side of Eq. (3.17a) approaches again $\pi N(\epsilon_F)/2$, with $N(\epsilon_F)$ the single–particle density of states at the Fermi level. However, since we are now dealing with interacting electrons, $N(\epsilon_F)$ includes interaction as well as disorder effects. $Y^{(1)}_{n_1 n_2}$ in that limit approaches some number $\pi Y^{(1)}/2$. We then obtain, for small $\Omega_{n_1-n_2}$,

$$X^{(1)}_{n_1 n_2} = \frac{\pi}{16|\Omega_{n_1-n_2}|}\left[N(\epsilon_F) + N_F\,(N_F\Gamma^{(s)})^2 Y^{(1)}\right] \quad.$$

$$(3.18)$$

Notice that the ψ-connected correlation functions vanish for noninteracting electrons, so that in an expansion in powers of the interaction constant, $Y^{(1)}$ is of $O(1)$. The term in brackets on the right–hand side of Eq. (3.18) is equal to $\pi\left(N(\epsilon_F)\right)^2/8H$, with H the quasi-particle or specific heat density of states.[5,6,16]

Similarly, one finds for the coefficients of $\delta_{n_1-n_3, n_2-n_4}$ in the Ward identity,

$$\Omega_{n_1-n_2}\, X^{(2)}_{n_1 n_2, n_3 n_4} = N_F(N_F\Gamma^{(s)})^2\, Y^{(2)}_{n_1 n_2, n_3 n_4}$$
$$-16\Gamma^{(s)}T\left\{ \left[G(\mathbf{x},\mathbf{x};\omega_{n_1}) - G(\mathbf{x},\mathbf{x};\omega_{n_2})\right] \right.$$
$$\left. \times \phi_{n_3 n_4}(\mathbf{x},\mathbf{x})\right\}_{dis} \quad, \qquad (3.19a)$$

where

$$\phi_{n_3 n_4}(\mathbf{x},\mathbf{x}) = \int d\mathbf{y}\, G(\mathbf{y},\mathbf{x};\omega_{n_3})\, G(\mathbf{x},\mathbf{y};\omega_{n_4}) \quad, \quad (3.19b)$$

and $Y^{(2)}$ is again given in terms of connected 4-ψ correlation functions. If desired, it can be calculated in perturbation theory in $\Gamma^{(s)}$. For $\Omega_{n_1 n_2} \to 0$, Eqs. (3.19) yield $X^{(2)}_{n_1 n_2, n_3 n_4} \sim T/|\Omega_{n_1-n_2}||\Omega_{n_3-n_4}|$. Again, the Gaussian propagator has that property, see Eqs. (2.37a), (2.39), (2.40).

From Eqs. (3.18) and (3.19) we see that $^0_0Q^{\alpha\alpha}_{nm}$ remains a soft mode in the presence of interactions. This is due to the frequency structure of the interaction term, which ensures that the 2-point Q-correlation function remains soft rather than acquiring a mass, as one might naively expect from an inspection of the action, Eqs. (2.21) - (2.24). Note that this softness is *not* restricted to the particular linear combination of Q that constitutes the particle number density. It therefore is not related to the particle number conservation law. Rather, it is a consequence of the spontaneous breaking of a continuous symmetry in the system, viz. the invariance, at zero external frequency, under the rotations between positive and negative frequencies discussed above, or between retarded and advanced Green functions. The soft modes in question are the corresponding Goldstone modes.

As in the noninteracting case, there is a variety of other soft modes that are related to the above Ward identity

by means of additional symmetries that are not broken. First of all, rotational symmetry in replica space again implies that the $^0_0Q^{\alpha\beta}_{nm}$, $nm < 0$, are also soft.[34] Next we show that the spin–triplet channel also remains soft. To this end, let us denote the (identical) coupling constants in the three branches of the spin–triplet channel by $\Gamma^{(t)}_i$, $i = 1, 2, 3$. Now we pick from the transformations T, Eq. (3.11a), the one that interchanges, say, $i = 0$ and $i = 1$. The action is still invariant under this transformation, provided that we also interchange $\Gamma^{(s)}$ and $\Gamma^{(t)}_1$. This shows that ^1_0Q also obeys Eq. (3.16), but with the $\Gamma^{(s)}$ in Eqs. (3.17) - (3.19a) replaced by $\Gamma^{(t)}$. Since this can be done for all branches of the spin–triplet channel, it follows that they all are soft modes. Alternatively, we can invoke rotational invariance in spin space (an $SU(2)$ subgroup of the large symplectic group), which implies that all three components of the spin–triplet have identical correlation functions. This argument also shows that there are no singlet–triplet cross correlations. By means of similar arguments one easily convinces oneself that the particle–particle channel is still soft, with the appropriate Cooper channel interaction amplitude replacing $\Gamma^{(s)}$ or $\Gamma^{(t)}$. For the action shown in Eq. (2.21) - (2.24), we thus conclude that all channels are still soft, and all the soft modes can be traced to the Ward identity, Eq. (3.14a), and additional symmetry relations. If some of the additional symmetries are broken externally, e.g. by means of an external magnetic field, or by magnetic impurities, then the respective modes acquire a mass, see Refs. 24,16.

4. Separation of soft and massive modes

From the previous subsection we know that the correlation functions of the Q_{nm} with $nm < 0$ are soft, while those with $nm > 0$ are massive. Our next goal is to separate these degrees of freedom in such a way that the soft modes remain manifestly soft to all orders in perturbation theory. To give a simple example, in an $O(N)$ ϕ^4–theory this is achieved by writing the N-component vector field as $\vec{\phi}(\mathbf{x}) = \rho(\mathbf{x})\,\hat{\phi}(\mathbf{x})$, with ρ the norm of the vector, and $|\hat{\phi}(\mathbf{x})| \equiv 1$. With this separation, only gradients of $\hat{\phi}$ appear in the theory, and so $\hat{\phi}$ is manifestly soft, see Ref. 35 and Appendix B. We are looking for an analogous separation of our matrix degrees of freedom. For noninteracting electrons this has been done in Refs. 17 and 18, and we will construct a suitable generalization of the procedure used by these authors.

The matrices Q that we are considering are complex $8Nn\times 8Nn$ matrices that obey Eqs. (2.26). Alternatively, we can consider the Q as $2Nn \times 2Nn$ matrices whose elements are spin–quaternions, or $4Nn \times 4Nn$ matrices whose elements are quaternions. For our purposes the last choice will be most convenient. We thus study a set of matrices Q with quaternion valued matrix elements $Q^{\alpha\beta}_{nm,\sigma\sigma'}$ ($n,m = 1,\ldots,N$; $\alpha,\beta = 1,\ldots,n$; $\sigma,\sigma' = \uparrow$

, ↓),[36]

$$Q_{nm,\sigma\sigma'}^{\alpha\beta} = \sum_{r=0}^{3} {}_r Q_{nm,\sigma\sigma'}^{\alpha\beta} \ \tau_r \quad , \qquad (3.20a)$$

that obey

$$Q = C^T Q^T C \quad , \qquad (3.20b)$$
$$Q^\dagger = -\Gamma Q \Gamma^{-1} \quad , \qquad (3.20c)$$

with Γ from Eq. (2.26c), and $C = \mathbb{1} \otimes \tau_2$, where $\mathbb{1}$ denotes the $4Nn \times 4Nn$ unit matrix. This set of matrices is invariant under transformations T that obey $T C T^T = C$, i.e. under symplectic transformations.[37]

For what follows, the constraint expressed by Eq. (3.20c) is somewhat awkward to handle. We therefore analytically continue the matrix elements of Q, considered as functions of the imaginary frequencies, $i\omega_n$ and $i\omega_m$, to the complex plane: $i\omega_n \to z_1$, and $i\omega_m \to z_2$. As functions of z_1 or z_2, the Q have a branch cut on the real axis, and from Sec. III B 3 we know that those matrix elements with z_1 and z_2 approaching the real axis from opposite sides are soft modes. The analytic continuation of Eq. (3.20c) reads

$$(Q^\dagger)_{z_1 z_2} = -Q_{z_1^* z_2^*} \quad . \qquad (3.21a)$$

In particular, for real frequencies, ω and ω', we have

$$(Q^\dagger)_{\omega \pm i0, \omega' \pm i0} = -Q_{\omega \mp i0, \omega' \mp i0} \quad . \qquad (3.21b)$$

If we analytically continue Q onto the unphysical Riemann sheet, $Q_{\omega,\omega'}$ is thus formally anti–hermitian. It therefore has $4Nn$ imaginary eigenvalues $\lambda_j \tau_0$ ($\lambda_j \in i\mathcal{R}$; $j = 1, \ldots, 4Nn$), and can be diagonalized by means of unitary transformations $\tilde{\mathcal{S}} \in \mathrm{U}(4Nn, \mathcal{Q})$ (i.e. by unitary $4Nn \times 4Nn$ matrices whose elements are quaternions). Now $\mathrm{U}(4Nn, \mathcal{Q})$ is isomorphic to the unitary symplectic group $\mathrm{USp}(8Nn, \mathcal{C}) \equiv \mathrm{U}(8Nn, \mathcal{C}) \cap \mathrm{Sp}(8Nn, \mathcal{C})$.[31] This means that the Q can be diagonalized by means of unitary matrices that are also symplectic, and hence leave the set of Q invariant. That is, the most general Q can be written

$$Q = \tilde{\mathcal{S}} \ D \ \tilde{\mathcal{S}}^{-1} \quad , \qquad (3.22)$$

where D is diagonal, and $\tilde{\mathcal{S}} \in \mathrm{USp}(8Nn, \mathcal{C})$.

However, diagonalization is more than we want. Since we know that the Q_{nm} with $nm < 0$ are soft, while those with $nm > 0$ are massive, we are interested in generating the most general Q from a matrix P that is block–diagonal in Matsubara frequency space,

$$P = \begin{pmatrix} P^> & 0 \\ 0 & P^< \end{pmatrix} \quad , \qquad (3.23)$$

where $P^>$ and $P^<$ are matrices with elements P_{nm} where $n, m > 0$ and $n, m < 0$, respectively. This can easily be achieved. Since the analytic continuations of $P^>$ and $P^<$

are anti–hermitian, the most general P can be obtained from D by an element U of $\mathrm{USp}(4Nn, \mathcal{C}) \times \mathrm{USp}(4Nn, \mathcal{C})$. The most general Q can therefore be written

$$Q = \mathcal{S} \ P \ \mathcal{S}^{-1} \quad , \qquad (3.24)$$

with $\mathcal{S} = \tilde{\mathcal{S}} U^{-1}$. The set of transformations $\mathcal{S}$ is the set of all cosets of $\mathrm{USp}(8Nn, \mathcal{C})$ with respect to $\mathrm{USp}(4Nn, \mathcal{C}) \times \mathrm{USp}(4Nn, \mathcal{C})$, i.e. the $\mathcal{S}$ form the homogeneous space $\mathrm{USp}(8Nn, \mathcal{C}) / \mathrm{USp}(4Nn, \mathcal{C}) \times \mathrm{USp}(4Nn, \mathcal{C})$. The corresponding most general Q on the imaginary frequency axis is generated from the most general P by a set of transformations that is isomorphic to this coset space, and that can be explicitly constructed by reversing the Wick rotation that led us from imaginary frequencies to real ones. For our purposes, we will not need this explicit construction, and we will not distinguish between the two isomorphic spaces.

This achieves the desired separation of our degrees of freedom into soft and massive ones. The massive degrees of freedom are represented by the matrix P, while the soft ones are represented by the transformations $\mathcal{S} \in \mathrm{USp}(8Nn, \mathcal{C}) / \mathrm{USp}(4Nn, \mathcal{C}) \times \mathrm{USp}(4Nn, \mathcal{C})$. To come back to the $O(N)$ example at the beginning of this subsection, the analogy is as follows. For the N-component vector, the direction $\hat{\phi}$ is a point on the $(N-1)$-sphere, which is isomorphic to the homogeneous space $O(N)/O(N-1) \times O(1)$, see Appendix B. The unitary–symplectic coset space identified above is a matrix generalization of this.

In order to formulate the field theory in terms of the soft and massive modes, one also needs the invariant measure $I[P]$, or the Jacobian of the transformation from the Q to the P and the $\mathcal{S}$, defined by

$$\int D[Q] \ldots = \int D[P] \, I[P] \int D[\mathcal{S}] \ldots \quad . \qquad (3.25)$$

Since we will not need to know the explicit form of $I[P]$ in what follows, we relegate its derivation to Appendix C.

5. A Ward identity based on a local symmetry

In addition to the Ward identity derived on the basis of a global symmetry in Secs. III A 2 and III A 3 above, for the derivation of an effective field theory in the next subsection we will also need an identity that is based on a local symmetry. Such relations with the structure of a Ward identity or a Noether theorem can be derived on the basis of either the replica or the gauge structure of the theory. For our present purposes, we consider a local $U(1)$ gauge transformation of the Grassmannian field theory, Eqs. (2.2), that is independent of imaginary time τ,

$$\psi_\sigma(x) \to e^{-i\alpha(\mathbf{x})} \psi_\sigma(x) \quad . \qquad (3.26)$$

We define a $(d+1)$-component current vector $q_\mu = (q_0, \mathbf{q})$ with 'spatial' components q_i $(i = 1, \ldots d)$ as

$$q_0(x) = \sum_\sigma \bar{\psi}_\sigma(x) \, \psi_\sigma(x) \quad , \qquad (3.27a)$$

$$\mathbf{q}(x) = \frac{-1}{2m} \sum_\sigma \left(\psi_\sigma(x) \, \nabla \bar{\psi}_\sigma(x) + \bar{\psi}_\sigma(x) \, \nabla \psi_\sigma(x) \right) \quad , \qquad (3.27b)$$

and add to the action a term S_A that describes a coupling of the current to a τ-independent vector potential $A_\mu = (A_0, \mathbf{A})$,

$$S_A = -i \int dx \, A_\mu(\mathbf{x}) \, q^\mu(x) - \frac{1}{2m} \int dx \, \mathbf{A}^2(\mathbf{x}) \, q_0(x) \quad . \qquad (3.28)$$

It is then straightforward to establish that the partition function is gauge invariant, i.e. it is invariant under the local gauge tranformation given by Eq. (3.26), supplemented by a transformation of the vector potential,

$$A_\mu(\mathbf{x}) \rightarrow A_\mu(\mathbf{x}) + \partial_\mu \alpha(\mathbf{x}) \quad , \qquad (3.29)$$

with $\partial_\mu = (\partial_\tau, \nabla)$. The invariance statement is

$$Z[A_\mu] = Z[A_\mu + \partial_\mu \alpha] \quad , \qquad (3.30)$$

A Taylor expansion of Eq. (3.30) immediately leads to the conservation law

$$\partial_\mu \left(\langle q_\mu(x) \rangle - \frac{i}{m} A_\mu(x) \langle q_0(x) \rangle \right) = 0 \quad , \qquad (3.31)$$

where $\langle \ldots \rangle$ denotes an average with respect to the action $S + S_A$. By putting $A_\mu = 0$, one simply obtains the continuity equation for the fermion number density. Another useful identity can be obtained by integrating Eq. (3.31) over space and time, differentiating with respect to $A_i(\mathbf{y})$, and then setting $A_\mu = 0$. The result is,

$$\frac{-1}{m} \sum_{\substack{n,m \\ \sigma, \sigma'}} \int d\mathbf{y} \left\langle \frac{\partial \bar{\psi}_{n,\sigma}(\mathbf{x})}{\partial x^i} \psi_{n,\sigma}(\mathbf{x}) \, \frac{\partial \bar{\psi}_{m,\sigma'}(\mathbf{y})}{\partial y^j} \psi_{m,\sigma'}(\mathbf{y}) \right\rangle$$
$$= \delta_{ij} \sum_{n,\sigma} \langle \bar{\psi}_{n,\sigma}(\mathbf{x}) \, \psi_{n,\sigma}(\mathbf{x}) \rangle \quad . \qquad (3.32)$$

This identity holds for a particular realization of the disorder. If we perform the ensemble average it still holds, with the brackets now including the disorder average in addition to the quantum mechanical expectation value. One then has a relation between a homogeneous current–current correlation function and the particle number density that is known as the f-sum rule. The above derivation makes it clear that it is closely related to the particle number conservation law. Within models or approximations that describe an effective single-particle problem,

the quantum mechanical average on the left–hand side of Eq. (3.32) factorizes. Furthermore, Eq. (3.32) then holds for each term in the n-summation separately, since in a single-particle problem there is no frequency mixing. Within approximations where the disorder average factorizes as well, the identity can be written in terms of disorder averaged one–particle Green functions as,

$$\delta_{ij} G(\mathbf{x} = 0, \omega_n) = \frac{1}{m} \int d\mathbf{y} \, [\partial_i G(\mathbf{x} - \mathbf{y}, \omega_n)]$$
$$\times [\partial_j G(\mathbf{y} - \mathbf{x}, \omega_n)] \quad , \qquad (3.33a)$$

or, in Fourier space,

$$\delta_{ij} \sum_{\mathbf{q}} G(\mathbf{q}, \omega_n) = \frac{-1}{m} \sum_{\mathbf{q}} q_i \, q_j \left(G(\mathbf{q}, \omega_n) \right)^2 \quad . \qquad (3.33b)$$

This identity will be useful later.

B. Effective field theory for disordered fermions

We are now in a position to formulate the theory in such a way that the soft and massive modes, respectively, are separated, and that the former remain manifestly soft to all orders in perturbation theory. This formulation also provides a technically satisfactory derivation of the generalized nonlinear sigma–model for interacting, disordered fermions. We will first discuss this derivation, as well as corrections to the sigma–model. Then we will discuss the Fermi–liquid FP and the critical FP that describes the Anderson–Mott metal–insulator transition, and show that the sigma–model provides an adequate description of either one.

1. The nonlinear sigma–model

We return to Eqs. (2.20), (2.21), and use the representation, Eq. (3.24), of Q in terms of $\mathcal{S}$ and P. It is convenient to define a transformed ghost field Λ by

$$\Lambda(\mathbf{x}) = \mathcal{S}^{-1}(\mathbf{x}) \, \tilde{\Lambda}(\mathbf{x}) \, \mathcal{S}(\mathbf{x}) \quad . \qquad (3.34)$$

In terms of these variables, the partition function for the replicated theory, Eq. (2.20), reads

$$\tilde{Z} = \int D[P] \, D[\mathcal{S}] \, D[\Lambda] \, \exp\left(\mathcal{A}[\mathcal{S}, P, \Lambda] + \mathrm{Tr} \ln I[P] \right) \quad . \qquad (3.35)$$

Here $I[P]$ is the invariant measure, Eq. (3.25), and the action $\mathcal{A}$ reads,

$$\mathcal{A}[\mathcal{S}, P, \Lambda] = \mathcal{A}_{dis}[P] + \mathcal{A}_{int}[\mathcal{S}P\mathcal{S}^{-1}]$$
$$+ \frac{1}{2} \mathrm{Tr} \ln \left(G_0^{-1} - i\mathcal{S}\Lambda\mathcal{S}^{-1} \right)$$
$$+ \int d\mathbf{x} \, \mathrm{tr} \left(\Lambda(\mathbf{x}) P(\mathbf{x}) \right) \quad , \qquad (3.36)$$

with $\mathcal{A}_{dis}$ and $\mathcal{A}_{int}$ defined by Eqs. (2.23) and (2.24).

Now we expand $\mathcal{S}$, P, and Λ about their saddle point values. Alternatively, one could expand about the exact expectation values. However, the 'equation of state' for the latter is just the saddle point equation of state plus loop corrections, and for our purposes the difference is irrelevant. We therefore do not distinguish between saddle point values and expectation values, and denote both by $\langle\mathcal{S}\rangle$, $\langle P\rangle$, and $\langle\Lambda\rangle$, respectively. As a result, we do not distinguish either between the saddle point Green function as defined in Eq. (2.36b), and the Green function that contains the full expectation value of Λ as a self energy. From Eqs. (2.32) it follows that

$$\langle\mathcal{S}\rangle = \mathbb{1}\otimes\tau_0 \quad , \tag{3.37a}$$

$$\langle P\rangle_{12} = \delta_{12}\,(\tau_0\otimes s_0)\,\frac{i}{2V}\sum_{\mathbf{p}}\Big[i\omega_{n_1} - \mathbf{p}^2/2m + \mu$$
$$-i\langle\Lambda_{12}\rangle\Big]^{-1}$$
$$\equiv \frac{\pi}{4}\,N_F\,(\pi_{12} + i\gamma_{12}) \quad ,$$
$$\equiv \delta_{12}\,(\tau_0\otimes s_0)\,\frac{\pi}{4}\,N_F\,(\pi_1 + i\gamma_1) \quad , \tag{3.37b}$$

$$\langle\Lambda_{12}\rangle = \frac{-2}{\pi N_F\tau_{rel}}\,\langle P_{12}\rangle - 4\Gamma^{(s)}T\sum_n e^{i\omega_n 0}\langle P_{nn}^{\alpha\alpha}\rangle \quad . \tag{3.37c}$$

and we write

$$P = \langle P\rangle + \Delta P \quad , \quad \Lambda = \langle\Lambda\rangle + \Delta\Lambda \quad . \tag{3.37d}$$

Here we have defined two new matrix fields, π and γ, for later reference. Their diagonal elements, π_1 and $i\gamma_1$, are odd and even functions, respectively, of their Matasubara frequency arguments. For small frequencies,

$$\pi_1 = \mathrm{sgn}\,\omega_{n_1} \quad , \tag{3.37e}$$

while γ_1 is a real constant that depends on the momentum cutoff in Eq. (3.37b).

Now we use the cyclic property of the trace to write the Tr ln-term in the action, Eq. (3.36), in the form

$$\mathrm{Tr}\ln\big(G_0^{-1} - \mathcal{S}i\Lambda\mathcal{S}^{-1}\big) = \mathrm{Tr}\ln\big(\mathcal{S}^{-1}G_0^{-1}\mathcal{S} - i\Lambda\big)$$
$$= \mathrm{Tr}\ln\big(G_{sp}^{-1}\big) + \mathrm{Tr}\ln\bigg(1 + G_{sp}\mathcal{S}^{-1}\,(\partial_\tau\mathcal{S})$$
$$+ \frac{1}{m}\,G_{sp}\mathcal{S}^{-1}\,(\nabla\mathcal{S})\,\nabla + \frac{1}{2m}\,G_{sp}\mathcal{S}^{-1}\,(\nabla^2\mathcal{S})$$
$$- G_{sp}i(\Delta\Lambda)\bigg) \quad , \tag{3.38}$$

Notice that now the transformation matrix $\mathcal{S}$ always appears in conjuction with some derivative, while the fluctuations $\Delta\Lambda$ are massive.[38] We proceed to expand in

powers of the derivatives of $\mathcal{S}$, and in powers of $\Delta\Lambda$, in analogy to the procedure used in Ref. 18. The simplest term in this expansion is

$$\mathrm{Tr}\,\big(G_{sp}\mathcal{S}^{-1}(\partial_\tau\mathcal{S})\big) = \int d\mathbf{x}\,\mathrm{tr}\,\big(i\Omega\,\mathcal{S}(\mathbf{x})\,G_{sp}(\mathbf{x}=0)$$
$$\times\mathcal{S}^{-1}(\mathbf{x})\big)$$
$$= \frac{\pi N_F}{2}\int d\mathbf{x}\,\mathrm{tr}\,\big(\Omega\hat{Q}(\mathbf{x})\big) + O\big(\Omega^2\hat{Q}\big) \quad . \tag{3.39}$$

Here we have defined a frequency matrix,

$$\Omega_{12} = (\tau_0\otimes s_0)\,\delta_{12}\,\omega_{n_1} \quad , \tag{3.40a}$$

and a new field, $\hat{Q}(\mathbf{x}) = \mathcal{S}(\mathbf{x})\,(\pi + i\gamma)\,\mathcal{S}^{-1}(\mathbf{x})$. Since the contribution of γ_{12} to $\langle P\rangle_{12}$ is, in the low–frequency limit, just an imaginary constant, it leads only to a constant contribution to $\hat{Q}$, which does not contribute to the lowest order terms in the present expansion. We can therefore neglect it, and write

$$\hat{Q}(\mathbf{x}) = \mathcal{S}(\mathbf{x})\,\pi\,\mathcal{S}^{-1}(\mathbf{x}) \quad , \tag{3.40b}$$

with π from Eq. (3.37b). Notice that $\hat{Q}$ is hermitian, and has the properties $\hat{Q}^2 = 1$, and $\mathrm{tr}\,\hat{Q} = 0$. $\hat{Q}$ will turn out to be a convenient parametrization of the soft modes, in analogy to the unit vector $\hat{\phi}$ in ϕ^4–theory, see Appendix B.

Now consider the gradient terms in the expansion of Eq. (3.38). To this end, it is convenient to define the vector field,

$$\mathbf{s}(\mathbf{x}) = \mathcal{S}^{-1}(\mathbf{x})\,(\nabla\mathcal{S})(\mathbf{x}) \quad . \tag{3.41}$$

with matrix valued components $s^i(\mathbf{x})$, $(i = 1,\ldots d)$. The expansion now proceeds in powers of $\mathbf{s}$, or $\nabla\mathcal{S}$. The term linear in $\mathbf{s}$ vanishes by symmetry. To quadratic order in $\mathbf{s}$, both the next–to–last term under the Tr ln (after a partial integration) and the square of the preceding term contribute. The result is a contribution to the action,

$$\frac{1}{4m}\sum_{12}\frac{1}{V}\sum_{\mathbf{q}}\eta_{12,ij}(\mathbf{q})\int d\mathbf{x}\,s_{12}^i(\mathbf{x})\,s_{21}^j(\mathbf{x}) \quad , \tag{3.42a}$$

with

$$\eta_{12,ij}(\mathbf{q}) = \delta_{ij}\frac{1}{2}\,[G_{sp}(\mathbf{q},\omega_{n_1}) + G_{sp}(\mathbf{q},\omega_{n_2})]$$
$$+ \frac{1}{m}\,q_i\,q_j\,G_{sp}(\mathbf{q},\omega_{n_1})G_{sp}(\mathbf{q},\omega_{n_2}) \quad . \tag{3.42b}$$

The approximation that neglects all fluctuations of P and Λ implies a factorization of four–point functions into products of Green functions, and hence Eq. (3.33b) applies. We thus have[39]

$$\lim_{n_1,n_2\to 0}\frac{1}{V}\sum_{\mathbf{q}}\eta_{12,ij}(\mathbf{q}) = \frac{1}{2}\,(1 - \pi_1\pi_2)\,\delta_{ij}\,\eta \quad , \tag{3.43a}$$

18

where

$$\eta = \frac{1}{2V} \sum_{\mathbf{q}} \left(G_+(\mathbf{q}) + G_-(\mathbf{q}) \right)$$
$$+ \frac{1}{dmV} \sum_{\mathbf{q}} \mathbf{q}^2 \, G_+(\mathbf{q}) \, G_-(\mathbf{q}) \quad , \qquad (3.43b)$$

with $G_\pm = G_{sp}(\omega_n \to \pm 0)$. The structure of Eq. (3.43a) allows one to write the low–frequency limit of the term quadratic in $\mathbf{s}$ in terms of the matrix $\hat{Q}$ defined in Eq. (3.40b),

$$\sum_{12} \int d\mathbf{x} \, s_{12}^i(\mathbf{x}) \, s_{21}^j(\mathbf{x}) \frac{1}{V} \sum_{\mathbf{q}} \eta_{12,ij}(\mathbf{q})$$
$$= \frac{-1}{4} \eta \int d\mathbf{x} \, \text{tr} \left(\nabla \hat{Q}(\mathbf{x}) \right)^2 + O(\nabla^3, \Omega\nabla) \,. \quad (3.44)$$

Finally, we can use Eq. (3.33b) again to rewrite the coupling constant η in a more familiar form,

$$\eta = \frac{-1}{2dmV} \sum_{\mathbf{q}} \mathbf{q}^2 \left[G_+(\mathbf{q}) - G_-(\mathbf{q}) \right]^2 = \pi \, m \, \sigma_0 \quad . \quad (3.45)$$

One recognizes σ_0 as the conductivity in the self–consistent Born approximation with the interaction taken into account in Hartree–Fock approximation.

The remaining task is to deal with the fields P and Λ. Since the fluctuations of these fields are massive modes, in order to capture the leading effects of the soft modes in the system it suffices to integrate them out in any approximation that preserves the Ward identities derived in Sec. III A. The simplest way to do so is to integrate them out in saddle–point approximation, i.e. to neglect ΔP and $\Delta\Lambda$ everywhere. From Eqs. (3.36), (3.37) we see that then the entire action can be expressed in terms of $\hat{Q}$. Finally, we subtract from $\hat{Q}$ its saddle–point value $\langle \hat{Q} \rangle = \pi$, and define a matrix field $\tilde{Q} = \hat{Q} - \pi$. Replacing $\hat{Q}$ by $\tilde{Q}$ everywhere in the action leads just to an uninteresting constant contribution. This is obvious for all terms except for the $r = 0$ channel in the spin–singlet interaction term. There a term linear in $\hat{Q}$ seems to remain, but inspection shows that it is proportional to $\text{tr} \, \hat{Q} = 0$. We then obtain the following effectice action,

$$\mathcal{A}_{NL\sigma M} = \frac{-\pi}{16} \sigma_0 \int d\mathbf{x} \, \text{tr} \left(\nabla \tilde{Q}(\mathbf{x}) \right)^2$$
$$+ \frac{\pi N_F}{4} \int d\mathbf{x} \, \text{tr} \left(\Omega \, \tilde{Q}(\mathbf{x}) \right) + \mathcal{A}_{int}[\tilde{Q}] \quad ,$$
$$(3.46a)$$

in terms of the matrix field $\tilde{Q} = \hat{Q} - \pi$, where π is the diagonal matrix defined in Eq. (3.37b), and $\hat{Q}$ is subject to the constraints,

$$\hat{Q}^2(\mathbf{x}) \equiv \mathbb{1} \otimes \tau_0 \quad , \quad \hat{Q}^\dagger = \hat{Q} \quad , \quad \text{tr} \, \hat{Q}(\mathbf{x}) = 0 \quad .$$
$$(3.46b)$$

We note that the hermiticity of $\hat{Q}$ is a consequence of our having dropped its anti–hermitian part, since it does not contribute to the leading terms in the long–wavelength, low–frequency effective action. The effective action given by Eqs. (3.46) is the generalized nonlinear sigma–model that was first proposed by Finkel'stein[15] as a model for interacting disordered electrons, and whose properties have been studied in considerable detail, in particular with respect to the metal–insulator transition that is described by the model.[16,40] In the next subsection, we will discuss $\mathcal{A}_{NL\sigma M}$, as well as its corrections, from a RG point of view, and show that it suffices for a correct description of both the critical FP and the stable Fermi–liquid FP, as well as the leading corrections to scaling at the latter.

2. The disordered Fermi–liquid fixed point

In this subsection we discuss the RG properties of the model derived in the previous section and, in particular, introduce the concept of a disordered Fermi–liquid FP. Physically, this FP characterizes a system with a finite density of states at the Fermi surface and diffusive two–particle excitations.

a. The fixed point action   To proceed, it is convenient to parametrize the matrix field $\hat{Q}$ in a way analogous to the π-field parametrization of the $O(N)$ vector model, see Appendix B. We thus write $\hat{Q}$ in a block matrix form analogous to that used in Eq. (3.23) as

$$\hat{Q} = \begin{pmatrix} \sqrt{1 - qq^\dagger} & q \\ q^\dagger & -\sqrt{1 - q^\dagger q} \end{pmatrix} \quad , \qquad (3.47)$$

where the matrix q has elements q_{nm} whose frequency labels are restricted to $n \geq 0$, $m < 0$. This representation builds in the constraints given by Eq. (3.46b). While the sigma–model can be entirely expressed in terms of $\hat{Q}$, this is not the case for the corrections to it. We therefore also express $\mathcal{S}$ in terms of q. For what follows, we will need only the first terms of an expansion in powers of q. From Eqs. (3.40b) and (3.47) we obtain,

$$\mathcal{S} = \mathbb{1} \otimes \tau_0 + \Delta\mathcal{S}$$
$$= \mathbb{1} \otimes \tau_0 + \frac{1}{2} \begin{pmatrix} 0 & -q \\ q^\dagger & 0 \end{pmatrix} + O(q^2) \quad . \qquad (3.48)$$

Now we perform a momentum-shell RG procedure.[41,42] For the rescaling part of this transformation, we need to assign scale dimensions to the soft field q, and to the massive fields ΔP, $\Delta\Lambda$ from Eq. (3.37d). Choosing the scale dimension of a length L to be $[L] = -1$, we write in analogy to Eqs. (B6),

$$[q(\mathbf{x})] = \frac{1}{2} (d - 2 + \eta') \quad , \qquad (3.49a)$$

$$[\Delta P(\mathbf{x})] = [\Delta\Lambda(\mathbf{x})] = \frac{1}{2} (d - 2 + \eta) \quad , \qquad (3.49b)$$

19

which defines the exponents η and η'. The stable Fermi–liquid FP of the theory is characterized by the choice,[43]

$$\eta = 2 \quad , \quad \eta' = 0 \quad . \qquad (3.50a)$$

Physically, $\eta' = 0$ corresponds to diffusive correlations of the q, and $\eta = 2$ means that the correlations of the ΔP are of short range. This is indeed what one expects in a disordered Fermi liquid. In addition, we must specify the scale dimension of frequency or temperature, i.e. the dynamical scaling exponent $z = [\omega] = [T]$. In order for the FP to be consistent with diffusion, that is with frequencies that scale like the squares of wave numbers, we must choose

$$z = 2 \quad . \qquad (3.50b)$$

Now we expand the sigma–model action, Eq. (3.46a), in powers of q. In a symbolic notation that leaves out everything not needed for power counting purposes, we write

$$\mathcal{A}_{NL\sigma M} = \frac{-1}{G} \int d\mathbf{x} \, (\nabla q)^2 + H \int d\mathbf{x} \, \omega \, q^2$$
$$+ \Gamma T \int d\mathbf{x} \, q^2 + O(\nabla^2 q^4, \omega q^4, T q^3) \quad , \quad (3.51)$$

with the bare coupling constants $G \sim 1/\sigma_0$, and $H \sim N_F$. Γ can stand for any of the three interaction coupling constants. Power counting shows that, with the above choices for the exponents, all of these coupling constants have vanishing scale dimensions with respect to our FP,

$$[G] = [H] = [\Gamma] = 0 \quad . \qquad (3.52)$$

These terms therefore make up part of the FP action.

Now consider first the corrections that arise within the sigma–model. The leading ones have been indicated in Eq. (3.51). We denote the corresponding coupling constants by $c_{\nabla^2 q^4}$, etc., with a subscript that identifies the structure of the respective contribution to the action. One finds

$$\left[c_{\nabla^2 q^4}\right] = \left[c_{\omega q^4}\right] = -(d-2) \quad . \qquad (3.53a)$$

With respect to the remaining term we note that it is of order q^3. Any contribution to physical q-correlation functions therefore contain this term squared, and therefore the relevant scale dimension is

$$\left[\left(c_{Tq^3}\right)^2\right] = 2\left[c_{Tq^3}\right] = -(d-2) \quad . \qquad (3.53b)$$

We see that all of these operators are irrelevant with respect to the Fermi–liquid FP for all dimensions $d > 2$, and that they become marginal in $d = 2$ and relevant in $d < 2$. All other terms that are contained in the sigma–model are more irrelevant than the ones considered.

We next consider corrections to the sigma–model action. From Eq. (3.36) we see that there are five classes

of such terms: (1) $\mathcal{A}_{dis}$, (2) the term tr (ΛP), (3) those parts of the Tr ln-term that contain the massive fluctuation $\Delta\Lambda$, (4) the parts of $\mathcal{A}_{int}$ that contain the massive fluctuation ΔP, and (5) contributions to $\langle P \rangle$ that were neglected in writing Eqs. (3.39) and (3.40b). Since all terms linear in the fluctuations vanish, class (1) just contributes

$$\mathcal{A}_{dis}[\Delta P] = -\frac{1}{\tau_1} \int d\mathbf{x} \, \left(\text{tr } \Delta P(\mathbf{x})\right)^2$$
$$+ \frac{1}{\tau_{rel}} \int d\mathbf{x} \, \text{tr } \left(\Delta P(\mathbf{x})\right)^2 \quad , \qquad (3.54)$$

where we have scaled the relaxation times τ_1 and τ_{rel} from Eqs. (2.23) with $1/2\pi N_F$ and $1/\pi N_F$, respectively. Power counting yields

$$[\tau_1] = [\tau_{rel}] = 0 \quad , \qquad (3.55)$$

so $\mathcal{A}_{dis}$ is part of the FP action. Class (2) also contributes a marginal term to the action, and so does the leading contribution of class (3), which has the structure Tr $(G\,\Delta\Lambda)^2$. Combining these two marginal terms, we have a contribution to the FP action,

$$\mathcal{A}_{\Lambda P} = \int d\mathbf{x} \, \text{tr } \left(\Delta\Lambda(\mathbf{x})\,\Delta P(\mathbf{x})\right)$$
$$+ \frac{1}{4} \int d\mathbf{x} \, d\mathbf{y} \, \text{tr } \left(G(\mathbf{x}-\mathbf{y})\,\Delta\Lambda(\mathbf{y})\,G(\mathbf{y}-\mathbf{x})\,\Delta\Lambda(\mathbf{x})\right),$$
$$(3.56)$$

apart from an uninteresting constant. The remaining terms of class (3) are irrelevant contributions that couple q and $\Delta\Lambda$. The least irrelevant terms have the structures Tr $(\Delta\Lambda \, \omega \, q)$, and Tr $(\Delta\Lambda \, \nabla^2 q)$. For contributions to physical correlation functions we must again consider the squares of these structures, and the relevant scale dimensions are

$$\left[\left(c_{\Delta\Lambda \, \omega \, q}\right)^2\right] = \left[\left(c_{\Delta\Lambda \, \nabla^2 \, q}\right)^2\right] = -2 \quad . \qquad (3.57)$$

These terms are more irrelevant than the corrections contained in the sigma–model for all $d < 4$. Moreover, due to the even integer value of the scale dimension, they lead only to analytic, and therefore uninteresting, corrections to scaling near the FP. The leading term of class (4) is one that couples q (from one of the $\mathcal{S}$) and ΔP. The corresponding coupling constant has a scale dimension

$$\left[\left(c_{Tq\,\Delta P}\right)^2\right] = -2 \quad . \qquad (3.58)$$

Again, these operators lead only to analytic corrections to scaling. Cross correlations between these various terms also appear, and the corresponding operators have the same scale dimension, namely -2, as the ones shown. Finally, it is easily checked that the least irrelevant terms of class (5) are of the structure $\omega^2 q^2$, and they also have a scale dimension

$$\left[c_{\omega^2 q^2}\right] = -2 \quad . \tag{3.59}$$

The above RG arguments show that the theory contains a Fermi–liquid FP that is stable for all $d > 2$, and analogous to the low–temperature FP of an $O(N)$ Heisenberg model. The FP action consists of the nonlinear sigma–model action $\mathcal{A}_{NL\sigma M}$, Eq. (3.46a), expanded to second order in q, and of $\mathcal{A}_{dis}$ and $\mathcal{A}_{\Lambda P}$, Eqs. (3.54) and (3.56). We have further shown that the leading nonanalytic corrections to scaling near the Fermi–liquid FP are contained in the nonlinear sigma–model. An effective action that describes both the FP and the leading corrections to scaling is therefore

$$\mathcal{A}_{eff} = \mathcal{A}_{NL\sigma M}[q] + \mathcal{A}_{dis}[\Delta P] + \mathcal{A}_{\Lambda P}[\Delta \Lambda, \Delta P] \quad . \tag{3.60}$$

Notice that the soft and massive modes do not couple in this effective action. Furthermore, the massive fields ΔP and $\Delta \Lambda$ appear only quadratically. If desired, then the latter can be integrated out to obtain an action entirely in terms of the physical fields q and ΔP.

b. Scaling behavior of observables We now discuss the physical meaning of the corrections to scaling induced by the irrelevant operators that we have identified above. Let us denote by the generic name u any of the least irrelevant operators whose scale dimension is $[u] = -(d-2)$, and let us discuss various observables, viz. the conductivity σ, the specific heat coefficient γ_V, the single–particle density of states N, and the spin susceptibility χ_s. Which of the various operators with scale dimension $-(d-2)$ is the relevant one depends on the quantity under consideration.

Let us first consider the dynamical conductivity, $\sigma(\omega)$. Its bare value is proportional to $1/G$, and according to Eq. (3.52) its scale dimension is zero. We therefore have the scaling law,

$$\sigma(\omega, u) = \sigma(\omega\, b^z, u b^{-(d-2)}) \quad , \tag{3.61a}$$

where b is an arbitrary RG scale factor. By putting $b = 1/\omega^{1/z}$, and using $z = 2$, Eq. (3.50b), as well as the fact that $\sigma(1, x)$ is an analytic function of x, we find that the conductivity has a singularity at zero frequency, or a long–time tail, of the form

$$\sigma(\omega) = \text{const.} + \omega^{(d-2)/2} \quad . \tag{3.61b}$$

This nonanalyticity is well known from perturbation theory for both noninteracting[44] and interacting[45] electrons. This shows that in this case either one of the coupling constants in Eqs. (3.53a) and (3.53b) can play the role of u. The present analysis proves that the $\omega^{(d-2)/2}$ is the exact leading nonanalytic behavior.

The specific heat coefficient, $\gamma_V = C_V/T$, is proportional to the quasiparticle density of states H,[5,6,46] whose scale dimension vanishes according to Eq. (3.52). We thus have a scaling law

$$\gamma_V(T, u) = \gamma_V(Tb^z, u b^{-(d-2)}) \quad , \tag{3.62a}$$

which leads to a low–temperature behavior

$$\gamma_V(T) = \text{const.} + T^{(d-2)/2} \quad . \tag{3.62b}$$

From perturbation theory it is known[45] that γ_V shows this behavior only for interacting electrons, while for noninteracting systems the prefactor of the nonanalyticity vanishes. In this case, therefore, u must be equated with $(c_{Tq^3})^2$. This can not be seen by our simple counting arguments.

The single–particle density of states, N, is proportional to the expectation value of Q, and to study the leading correction to the finite FP value of N it suffices to replace Q by $\hat{Q}$. Then we have, in symbolic notation, $N \sim 1 + \langle q\, q^\dagger \rangle + \ldots = 1 + \Delta N$. The scale dimension of ΔN is $[\Delta N] = 2[q] = d - 2$. We find the scaling law

$$\Delta N(\omega) = b^{-(d-2)}\, \Delta N(\omega b^z) \quad , \tag{3.63a}$$

which leads to the so–called Coulomb anomaly,[47]

$$N(\omega) = \text{const.} + \omega^{(d-2)/2} \quad , \tag{3.63b}$$

Again, this behavior is known to occur only in the presence of electron–electron interactions.

Finally, we consider the wave vector dependent spin susceptibility, $\chi_s(\mathbf{q})$. χ_s is given by a Q-Q correlation function, and the leading correction to the finite Fermi–liquid value is obtained by replacing both of the Q by q. Then we have a term of the structure $\chi_s \sim T \int d\mathbf{x}\, \langle q^\dagger q \rangle$, with scale dimension $[\chi_s] = 0$. The relevant scaling law is

$$\chi_s(\mathbf{q}, u) = \chi_s(\mathbf{q}b, u b^{-(d-2)}) \quad , \tag{3.64a}$$

which leads to a nonanalytic dependence on the wave number,

$$\chi_s(\mathbf{q}) = \text{const.} + |\mathbf{q}|^{(d-2)} \quad . \tag{3.64b}$$

This behavior is also known from perturbation theory, and holds only for interacting electrons. It has recently been shown to have interesting consequences for the theory of ferromagnetism.[19]

To summarize, we see from the above arguments that all of the so–called weak–localization effects,[48] i.e. nonanalytic dependencies of various observables on frequency, temperature, or wave number in disordered electron systems that are well known from perturbation theory, emerge naturally in the present context as the leading corrections to scaling near the Fermi–liquid FP of a general field theory for disordered interacting electrons. Apart from providing an aesthetic, unifying, and very simple explanation for these effects, our arguments also prove that they do indeed constitute the leading nonanalytic behavior, a conclusion that cannot be drawn from perturbation theory alone. We emphasize the ease with which these results are obtained within the present

framework. This reflects the judicious choice of our starting point for the RG analysis. The transformation from fermionic variables to composite ones, together with the symmetry analysis that identifies the soft modes, provides for a formulation of the theory that allows for rather nontrivial results to be obtained from a simple RG analysis at the Gaussian level. Similar benefits can be obtained from analogous treatments of other systems, see Appendix B.

Finally, it is interesting to note that similar nonanalyticities occur in very different systems, for example, in classical fluids. In that context they are known as long–time tail effects, and were first discussed theoretically by using many body perturbation theory and mode coupling theory.[49] Later, they were examined using RG ideas, and they were shown to be related to corrections to the scaling behavior near a hydrodynamic FP.[50]

3. The critical fixed point

In addition to the stable Fermi–liquid FP discussed above, the generalized matrix nonlinear sigma–model given by Eq. (3.46a) also possesses another FP that in general has the properties of a critical FP. It characterizes a zero–temperature metal–insulator transition in dimensions greater than two. For a review of this FP, and of the various universality classes for the metal–insulator transition, we refer the reader elsewhere.[16] Here we briefly comment on the differences between the Fermi–liquid FP and the critical one, and on the justification for using the sigma–model to describe the critical behavior.

First of all, at the critical point the spontaneously broken symmetry is restored, so the distinction between positive and negative frequencies vanishes as the magnitude of the frequencies goes to zero.[51] Physically this corresponds to a vanishing single–particle of density of states at the Fermi level, which plays the role of an order parameter. This characterizes the metal–insulator transition as being of Anderson–Mott type. For example, the left–hand side of Eq. (3.17a) vanishes at the transition as $n_1, n_2 \to 0$. Therefore, matrix elements $\hat{Q}_{n_1 n_2}$ with $n_1 n_2 > 0$ and $n_1 n_2 < 0$, respectively, should both scale in the same way. This implies that at the critical FP, η' in Eq. (3.49a) is given by

$$\eta' = 2 - d \equiv -\epsilon \quad , \tag{3.65}$$

so that q is dimensionless. The physical critical exponent, η, on the other hand, is given by

$$\eta = -\epsilon + 2\beta/\nu \quad . \tag{3.66}$$

To see this, we note that $\hat{Q}$ is proportional to $\langle P \rangle$, which in turn is proportional to the density of states which vanishes at the transition with a critical exponent β.[52] Here ν is the critical exponent that describes the divergence of the localization or correlation length ξ. It occurs in

the scaling equality, Eq. (3.66), since η characterizes momentum singularities at criticality.

We add some comments on the scaling of ΔP fluctuations. At the critical point, the ΔP fluctuations are actually of long range, and the sigma–model description breaks down. To avoid this problem, one must work at sufficiently small momenta p (or frequencies), and at sufficiently large distances t from the critical point, so that the q fluctuations still dominate over the ΔP fluctuations. Since the ΔP correlation function is related to the longitudinal susceptibility, it diverges as $t^{-\gamma}$ while the q correlations diverge as $p^{-2+\eta}$.[53] The relevant inequality is therefore

$$p \ll t^{\gamma/(2-\eta)} = t^{\nu} \quad , \tag{3.67}$$

where we have used the scaling equality $\nu = \gamma/(2 - \eta)$. This can also be seen more explicitly by considering the ΔP-ΔP propagator. From Eq. (3.54), and the corresponding term that is proportional to $\left(\nabla(\Delta P)\right)^2$ one sees that the ΔP-mass m_P has a scale dimension of 1 (see also Appendix B 3). The ΔP fluctuations can therefore be neglected for momenta that are small compared to $m_P \sim \xi^{-1} \sim t^{\nu}$, which is again Eq. (3.67). In this regime the ΔP fluctuations are effectively of short range, just as they are at the Fermi–liquid FP, and the sigma–model is valid. Even though the range of validity of the sigma–model goes to zero as the critical point is approached, it still can be used to extract the critical behavior. The salient point is that the RG, and the scaling behavior derived from it, allows one to extrapolate from the regime given by Eq. (3.67) to the critical region. However, no information can be obtained in this way about the symmetric phase, i.e. the insulator.

Apart from these theories that work near $d = 2$, the matrix nonlinear sigma–model has also been studied in high dimensions.[40] This work has established $d_c^+ = 6$ as an upper critical dimension, above which the Anderson–Mott transition is correctly described by a simple Landau–type theory. This treatment of the problem stresses that the metal–insulator transition problem is, somewhat counterintuitively, conceptually simpler in the presence of electron–electron interactions than in their absence, since the interacting problem possesses a simple critical order parameter, viz. the density of states at the Fermi level. Also, in these papers the presence of random–field like terms in the renormalized action was discovered. These terms are responsible for the upper critical dimension being 6 (rather than 4 as in the $O(N)$ vector model), and they have led to the suggestion that the Anderson–Mott transition may have features reminiscent of a glass transition.[54]

4. Ferromagnetic transition in the metallic phase

It has been known for some time that the matrix nonlinear sigma–model, Eq. (3.46a), contains another critical

FP that is not related to a metal–insulator transition.[55] While originally the nature of the corresponding phase transition was not clear, we have recently shown that it is the zero–temperature transition from a paramagnet to a ferromagnet in a disordered metal in dimensions greater than two.[19] In this reference, a Landau–Ginzburg–Wilson (LGW) functional for spin density, or order–parameter, fluctuations has been derived. To obtain this functional, we integrated out all excitations or modes other than the order parameter, including all of the soft modes related to weak localization effects, see Sec. III B 2 above. In the effective, or order–parameter, field theory these extra soft modes led to nonanalyticities in the bare LGW functional. Using renormalization group methods, we were able to exactly determine the critical behavior at the magnetic phase transition.[19] The earlier approach of Ref. 55, on the other hand, had focused entirely on the behavior of the diffusion modes across the magnetic phase transition, while the order parameter fluctuations had effectively been integrated out. In some respects, these two approaches to the problem were therefore complementary to one another.

Even though both of these theories led to the same results, their physical underpinnings are very different, and neither one of them constitutes what one might consider the physically most obvious approach. Physically, the most sensible procedure would be to treat all of the soft modes at this quantum phase transition on the same footing. These soft modes include the diffusive modes discussed above, i.e. the $Q_{n_1 n_2}$ with $n_1 n_2 < 0$, as well as the fluctuations of the magnetic order parameter, that is, the spin density or magnetization. In terms of Q-fields, the local spin density at frequency ω_n, $n_s^i(\mathbf{x}, \omega_n)$, is given by (cf. Eq. (2.12b)

$$n_s^i(\omega_n, \mathbf{x}) = \sqrt{T} \sum_m \sum_{r=0,3} (-1)^r \, {}_r^i Q_{m,m+n}^{\alpha\alpha}(\mathbf{x}) \quad , \quad (3.68)$$

and the macroscopic magnetization order parameter, $\mathbf{M}$, has components,

$$M^i = \left\langle n_s^i(\omega_n = 0, \mathbf{x}) \right\rangle$$
$$= \sqrt{T} \sum_m \sum_{r=0,3} (-1)^r \left\langle {}_r^i Q_{mm}^{\alpha\alpha}(\mathbf{x}) \right\rangle \quad . \quad (3.69)$$

From these equations we see that, as the magnetic transition is approached, the classification of soft modes that was given in Section III A above, breaks down. New soft modes related to the criticality of Q_{mm} occur. Alternatively, the extra soft mode that appears near the ferromagnetic transition can be related to a new zero eigenvalue in the Gaussian eigenvalue problem. The matrix M in Eqs. (2.37a) and (2.39a) is proportional to,

$$
{}_0^i M_{n_1 n_2, n_3 n_4}(p) \sim \delta_{n_1 n_3} \delta_{n_2 n_4} \mathcal{D}_{n_1 n_2}^{-1}(\mathbf{p})
$$
$$
- 2T\Gamma_t \, \delta_{n_1 - n_2, n_3 - n_4} \quad , \quad (3.70)
$$

with $i = 1, 2, 3$ and we have considered the term diagonal in all of the replica labels since these terms contain the

interactions that lead to magnetism. The corresponding eigenvalue problem is,

$$
\mathcal{D}_{n_1 n_2}^{-1}(\mathbf{p}) f_{n_1 n_2}^i(\mathbf{p}) - 2\Gamma_t T \sum_n f_{n, n-(n_1-n_2)}^i(\mathbf{p})
$$
$$
= \lambda f_{n_1 n_2}^i(\mathbf{p}) \quad , \quad (3.71)
$$

with f the eigenfunction and λ the eigenvalue. Setting $n_1 = n_2$ and integrating over n_1 leads to an equation for λ,

$$
1 = 2\Gamma_t T \sum_n \frac{\mathcal{D}_{nn}(\mathbf{p} = 0)}{1 - \lambda \mathcal{D}_{nn}(\mathbf{p} = 0)} \quad . \quad (3.72)
$$

Criticality is characterized by a new zero eigenvalue, $\lambda = 0$, and we have put $\mathbf{p} = 0$, since this is where the first zero eigenvalue occurs. Eq. (3.72) reduces to a Stoner criterion for the occurrence of ferromagnetism, modified by disorder. Expanding in powers of λ leads to

$$
\lambda \sim -t \quad , \quad (3.73)
$$

with t the dimensionless distance from the critical point. From Eq. (3.71), we see that the critical eigenfunction is

$$
f_{nn}^i(\mathbf{p}) = \mathcal{D}_{nn} 2\Gamma_t \sqrt{T} \, n_s^i(\omega_n = 0, \mathbf{p}) \quad . \quad (3.74)
$$

The above discussion makes it clear that a physically satisfactory theory of the ferromagnetic, or any other, phase transition in a disordered metal should take both the soft modes that were discussed in the preceding subsections, and the additional, critical, soft modes into account on the same footing. Technically one needs to extract the critical part from the P-fluctuations. The details of such a theory of the paramagnetic–to–ferromagnetic phase transition will be discussed elsewhere.[27] Here we just mention why the sigma–model, in Refs. 55, manages to describe the correct critical behavior, even though it neglects the soft modes related to the magnetization. The salient point is again that there is a regime, described by Eq. (3.67), where the diffusive modes that are contained in the sigma–model dominate over the magnetization fluctuations that are not. As in the case of the metal–insulator transition, the RG allows one to extract the critical behavior from an analysis of that region.

IV. THE CLEAN LIMIT

In this section we discuss the clean limit of our field theory. For the reasons that were pointed out after Eqs. (2.13) in Sec. II A, this treatment of clean Fermi systems will not be as complete as our theory for the disordered case. For notational simplicity we will also suppress the particle–particle degrees of freedom, but they can be easily restored. Even with our restriction to two Fermi–liquid parameters (in the absence of the Cooper channel) instead of infinitely many, we will be able to study fundamental structural properties of the clean limit that will have to be included in any more complete theory.

A. Effective Q-field theory

Let us return to the action, Eq. (2.21), and perform the clean limit, $\tau_1, \tau_{rel} \to \infty$. As pointed out earlier, the theory has been purposely set up so that this limit can be taken without difficulty. $\mathcal{A}_{dis}$ then vanishes. The saddle point solution, Sec. II C, reduces to the ordinary Hartree–Fock approximation, and the Gaussian approximation reduces to RPA. In particular, the saddle point action contains an exact description of the noninteracting electron gas. This will be important in what follows.

As was pointed out at the end of Section II, the fluctuations of $\bar{\Lambda}$ are not massive in the clean case, since the function φ, Eq. (2.38b), is singular in the long–wavelength, low–frequency limit, see Eq. (2.43). Indeed, Eqs. (2.37) and (2.39) show that the Gaussian $\bar{\Lambda}$-propagator is just minus the noninteracting part of the Q-propagator. On the general principle that one should keep all of the soft modes on the same footing, this seems to suggest keeping the $\bar{\Lambda}$ field at the same level as the Q-field. Dropping the constant saddle–point contribution, and expanding the Tr ln in Eq. (2.21), the action then reads

$$\mathcal{A}[Q, \bar{\Lambda}] = \mathcal{A}_G^{clean}[Q, \bar{\Lambda}] + \Delta\mathcal{A}[Q, \bar{\Lambda}] \quad , \qquad (4.1a)$$

where $\mathcal{A}_G^{clean} = \mathcal{A}_G - \mathcal{A}_{dis}$ with $\mathcal{A}_G$ and $\mathcal{A}_{dis}$ from Eq. (2.36a), and

$$\Delta\mathcal{A}[Q, \bar{\Lambda}] = -\sum_{M=3}^{\infty} \frac{2^{M-1}}{M} \frac{i^M}{V^{(M-2)/2}} \sum_{\mathbf{k}_1, \dots \mathbf{k}_M} \sum_{n_1 \dots n_M}$$
$$\times \delta_{\mathbf{k}_1 + \dots + \mathbf{k}_M, 0} \; \chi_{n_1 \dots n_M}^{(M)}(\mathbf{k}_1, \dots, \mathbf{k}_{M-1})$$
$$\times \varphi_{n_1 n_2}^{-1}(\mathbf{k}_1) \cdots \varphi_{n_M n_1}^{-1}(\mathbf{k}_M)$$
$$\times \text{tr} \left[\left((\delta\bar{\Lambda})_{n_1 n_2}(\mathbf{k}_1) - (\delta Q)_{n_1 n_2}(\mathbf{k}_1) \right) \cdots \right.$$
$$\left. \times \left((\delta\bar{\Lambda})_{n_M n_1}(\mathbf{k}_M) - (\delta Q)_{n_M n_1}(\mathbf{k}_M) \right) \right] \quad .$$
$$(4.1b)$$

Here φ is given by Eq. (2.38b), and

$$\chi_{n_1 \dots n_M}^{(M)}(\mathbf{k}_1, \dots, \mathbf{k}_{M-1}) = \frac{1}{V} \sum_{\mathbf{p}} G_{sp}(\mathbf{p}, \omega_{n_1})$$
$$\times G_{sp}(\mathbf{p} + \mathbf{k}_1, \omega_{n_2}) \cdots G_{sp}(\mathbf{p} + \mathbf{k}_{M-1}, \omega_{n_M}) \quad . \quad (4.1c)$$

In this formulation of the theory the vertices are given by the product of the M-point correlation function $\chi^{(M)}$ and the M factors of φ^{-1} in Eq. (4.1b). An inspection of Eq. (4.1c) shows that the generic behavior of $\chi^{(M)}$, considered as a function of some generic wave number k, is characterized by a divergence $1/k^{M-1}$ for small k. This is the most divergent behavior $\chi^{(M)}$ can display, and it is realized unless either all of the frequencies ω_{n_1} through ω_{n_M} are positive, or all of them are negative, in which case $\chi^{(M)}$ scales like a constant. The behavior of the product of the φ^{-1}, on the other hand, depends on the detailed distribution of the frequency labels. If $n_i n_{i+1} > 0$, then $\varphi_{n_i n_{i+1}}^{-1}$ is a number, while for

$n_i n_{i+1} < 0$ it goes like k for small k. If only two of the frequency pairs have elements with opposite signs (this is the smallest nonzero number possible), then the complete vertex scales like $1/k^{M-3}$. For $M > 3$, the vertices are thus not finite in the limit of long wavelengths and small frequencies. In other words, the Q-$\bar{\Lambda}$ field theory given by Eqs. (4.1) is not local. The reason for this is as follows. As was mentioned in Sec. II, in the clean limit the one–particle excitations are soft, as is manifested by the massless single–particle Green function. These soft modes have been integrated out when we integrated out the fermions, and this is what leads to the nonlocalities in the effective field theory. The massless single–particle excitations are also indirectly responsible for the softness of the $\bar{\Lambda}$ fluctuations, and for the need to keep infinitely many Landau parameters. All of these difficulties thus have the same underlying source.

Given that the matrix field theory is nonlocal in any case, we can proceed and integrate out $\bar{\Lambda}$, as it will turn out that this does not lead to further undesirable properties of the theory. The integrating out of $\bar{\Lambda}$ can be done exactly in the sense of a prescription for doing perturbation theory for the resulting Q-field theory. To see this, we recall that the $\bar{\Lambda}$ propagator is minus the noninteracting part of the Q-propagator. As a result, the integration over $\bar{\Lambda}$ just cancels the noninteracting parts of any internal Q-propagators in a loop expansion for any Q-correlation function. This is easy to see for the first few terms in the loop expansion, and it can be proved by induction to be true order by order in perturbation theory. As an illustrative example, the cancellation scheme is demonstrated diagrammatically in Figs. 2 and 3 for the two–point vertex function, and the two–point propagator, respectively, to one–loop order. Notice that, in order to avoid double counting, one effect of the $\bar{\Lambda}$-field *must* be the cancellation of the noninteracting contributions, since the saddle–point contribution to the action already contains a complete description of the noninteracing electron system. What is remarkable is that it does nothing else. We then obtain the following effective action entirely in terms of Q-fields,

$$\mathcal{A}[Q] = \mathcal{A}_0[Q] + \mathcal{A}_{int}[Q] + \Delta\mathcal{A}[Q] \quad , \qquad (4.2a)$$

where

$$\mathcal{A}_0[Q] = -4 \sum_{\mathbf{k}} \sum_{nm} \sum_{r,i} {}_r^i(\delta Q)_{nm}(\mathbf{k}) \, \varphi_{nm}^{-1}(\mathbf{k})$$
$$\times {}_r^i(\delta Q)_{nm}(-\mathbf{k}) \quad , \qquad (4.2b)$$

is the noninteracting part of the Gaussian action, and

$$\mathcal{A}_{int}[Q] = -8 \sum_{i=0}^{3} \sum_{r=0,3} \Gamma^{(i)} \sum_{\mathbf{k}} T \sum_{n_1, n_2, n_3, n_4} {}_r^i(\delta Q)_{n_1 n_2}(\mathbf{k})$$
$$\times \delta_{n_1 - n_2, n_3 - n_4} {}_r^i(\delta Q)_{n_3 n_4}(-\mathbf{k}) \quad , \qquad (4.2c)$$

is the interacting one. The $\Gamma^{(i)}$ ($i = 0, 1, 2, 3$) were defined after Eq. (2.39b). The non–Gaussian part of the action reads,
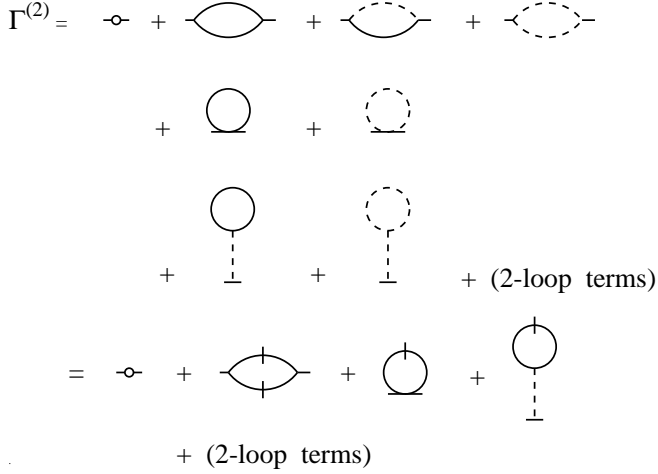
FIG. 2. Perturbation theory to one–loop order for the 2-point Q-vertex function. The small circle denotes the Gaussian vertex, the solid and dashed lines are Q-propagators and $\bar{\Lambda}$-propagators, respectively, and the solid lines with a vertical bar denote the interacting part of the Q-propagator.

$$\Delta\mathcal{A}[Q] = -\sum_{M=3}^{\infty} \frac{2^{M-1}}{M} \frac{(-i)^M}{V^{(M-2)/2}} \sum_{\mathbf{k}_1,\ldots\mathbf{k}_M} \sum_{n_1\ldots n_M}$$
$$\times\delta_{\mathbf{k}_1+\ldots+\mathbf{k}_M,0}\ \chi_{n_1\ldots n_M}^{(M)}(\mathbf{k}_1,\ldots,\mathbf{k}_{M-1})$$
$$\times\varphi_{n_1 n_2}^{-1}(\mathbf{k}_1)\cdots\varphi_{n_M n_1}^{-1}(\mathbf{k}_M)$$
$$\times\mathrm{tr}\left[(\delta Q)_{n_1 n_2}(\mathbf{k}_1)\cdots(\delta Q)_{n_M n_1}(\mathbf{k}_M)\right] \quad . \quad (4.2\mathrm{d})$$

This effective action needs to be supplemented by the following rules for doing perturbation theory:

rule 1: For calculating Q-propagators, all internal propagators must be taken as the interacting part of the Gaussian propagator, i.e. as the second term on the right–hand side of Eq. (2.39a).

rule 2: For calculating Q-vertex functions, rule 1 also applies. In addition, one needs to consider all reducible diagrams (which normally do not contribute to the vertices), with all reducible propagators replaced by minus the noninteracting Gaussian Q-propagator, i.e. minus the first term on the right–hand side of Eq. (2.39a).

Explicit calculations using these rules readily show that the integrals that correspond to the diagrams in a loop expansion are identical to integrals that one encounters in standard many–body perturbation theory for the same quantity. This allows for a one–to–one correspondence between many–body diagrams and the loop expansion based on the present field theory. Nevertheless, even at a calculational level the present formulation provides advantages compared to standard perturbation theory. For instance, due to the above rules, the loop expansion is equivalent to an expansion in powers of the screened Coulomb interaction, with the zeroth order, i.e. the Gaussian theory, reproducing RPA. The loop expansion therefore allows for a systematic improvement over RPA. Perhaps more importantly, our field–theoretical
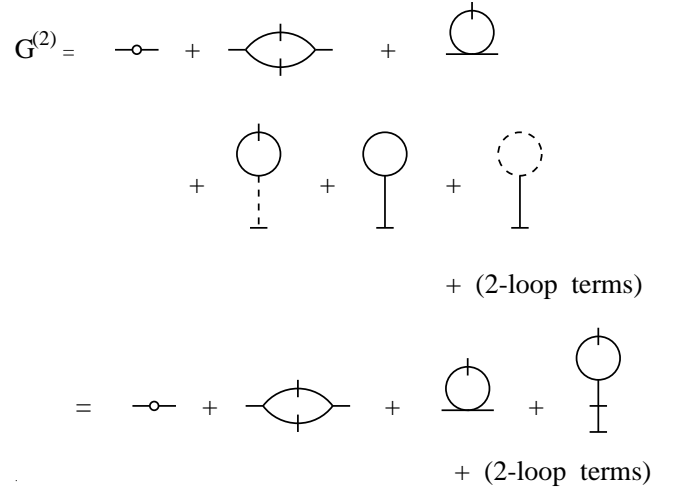

FIG. 3. Perturbation theory to one–loop order for the 2-point Q-propagator. The notation is the same as in Fig. 2.

formulation allows for an application of the renormalization group to draw structural conclusions about the theory in analogy to those discussed for the disordered case in Sec. III B 2. This we will discuss next.

B. The Fermi–liquid fixed point

We now are looking for a stable RG fixed point that describes the clean Fermi liquid, in analogy to the disordered Fermi–liquid FP of Sec. III B 2. For this purpose it is again convenient to split the matrix Q into blocks in frequency space,

$$Q_{nm} = \begin{cases} P_{nm} & \text{if } nm > 0 \\ q_{nm} & \text{otherwise} \end{cases} \quad . \quad (4.3)$$

As in Sec. III B 2, we define the scale dimension of a length L to be $[L] = -1$, and we define exponents η and η' by writing

$$[q(\mathbf{x})] = \frac{1}{2}\left(d-1+\eta'\right) \quad , \quad (4.4\mathrm{a})$$

$$[\Delta P(\mathbf{x})] = \frac{1}{2}\left(d-1+\eta\right) \quad . \quad (4.4\mathrm{b})$$

Here $\Delta P = P - \langle P\rangle$, and as in Sec. III we do not distinguish between ΔP and δP. The FP action has the properties one expects from a Fermi liquid if we choose

$$\eta = 1 \quad , \quad \eta' = 0 \quad , \quad (4.5\mathrm{a})$$

and a dynamical exponent

$$z = 1 \quad . \quad (4.5\mathrm{b})$$

Power counting shows that with these choices, $\mathcal{A}_0$ is dimensionless, and hence part of the FP action. So is the

part of $\mathcal{A}_{int}$ that is quadratic in q. The parts of $\mathcal{A}_{int}$ that couple q with ΔP, and ΔP with itself, are irrelevant with scale dimensions of $-1/2$ and -1, respectively.

Now consider the term of $O(Q^M)$ in the non–Gaussian part of the action, $\Delta\mathcal{A}$, Eq. (4.2d). Let N of the Q-fields be P's, and $M - N$ be q's. Denote the coupling constant for that term by $u_{N,M-N}$. Taking into account the properties of $\chi^{(M)}$ and φ discussed after Eqs. (4.1), we obtain

$$[u_{M,0}] = -(d+1)\frac{(M-2)}{2} \quad , \qquad (4.6a)$$

for $N = M$, and

$$[u_{N,M-N}] = -\frac{1}{2}(M-N-2) - (d-1)\frac{(M-2)}{2} \quad , \qquad (4.6b)$$

for $N \neq M$. Here we have made use of the following observation. Because of the two rules in the preceding subsection, all contractions that are performed in $\Delta\mathcal{A}$ to obtain renormalizations of the Gaussian action result in interacting propagators. Since the interaction carries a factor of temperature, see Eq. (4.2c), this results effectively in $(M-2)/2$ factors of T that need to be taken into account in addition to the terms that show explicitly in Eq. (4.2d). This factor of $T^{(M-2)/2}$ corresponds to the $M - 2$ internal Q-fields that get contracted in a renormalization of the Gaussian term. We note that this feature is automatically built into the theory only after the $\bar{\Lambda}$-field has been integrated out. Doing so is therefore advantageous. Since $M - N$ is necessarily even, we see from Eqs. (4.6) that all of the non–Gaussian terms are RG irrelevant with respect to the Fermi–liquid FP, provided that $d > 1$. For $d = 1$ there is an infinite set of marginal operators, which signals the instability of the Fermi–liquid ground state against the formation of a Luttinger liquid.[56] For $d < 1$ the Fermi liquid is unstable due to infinitely many relevant operators, as expected.

We mention that there is no consensus on whether or not interacting Fermi systems in $d > 1$ are necessarily Fermi liquids. In particular, Anderson[57] has proposed that there exists a stable FP that corresponds to a Luttinger–type liquid, at least in certain 2-d models. While we do not find such a FP, we stress that the above considerations do not constitute a proof that none exists. All we have shown is that the assumption of a Fermi–liquid FP in $d > 1$ leads self–consistently to the conclusion that this FP is stable. This is in agreement with a variety of other RG arguments,[8,58] and with explicit calculations.[59] However, we do not know what the basin of attraction for the Fermi–liquid FP is, and we cannot exclude the existence of other fixed points.

Obviously, the discussion of the corrections to scaling and their consequences for the behavior of thermodynamic quantities that was given for the disordered case in Sec. III B 2 can be carried over. The only difference is that the scale dimension of the least irrelevant operators,

which we again denote by u, is now $[u] = -(d-1)$. As an explicit example, let us consider the spin susceptibility, χ_s. It obeys a scaling law in analogy to Eq. (3.64a),

$$\chi_s(\mathbf{q}, u) = \chi_s(\mathbf{q}b, ub^{-(d-1)}) \quad , \qquad (4.7a)$$

which leads to a nonanalytic dependence on the wave number,

$$\chi_s(\mathbf{q}) = \text{const.} + |\mathbf{q}|^{d-1} \quad . \qquad (4.7b)$$

This is the leading wave number dependence of χ_s for $1 < d < 3$, and the leading nonanalytic one in all dimensions. In $d = 3$ there is a logarithmic correction to scaling, which our power counting arguments are not sensitive to, and the behavior is $\mathbf{q}^2 \ln|\mathbf{q}|$. This behavior has recently been obtained by means of explicit perturbative calculations,[60] and its implications for the paramagnet–to–ferromagnet transition at zero temperature have been discussed.[20] The nonanalytic behavior of other thermodynamic quantities, and of the quasiparticle inelastic life time, can be understood by means of analogous arguments.

V. CONCLUSION

In this paper we have given a general method to study the long–wavelength, low–frequency behavior of many–fermion systems, both with and without quenched disorder. The crucial ideas are to first identify the slow modes of the system by using a symmetry analysis, then to separate these soft modes from the massive ones, and, finally, to use renormalization group ideas to eliminate the degrees of freedom that are irrelevant in the long–wavelength limit.

Using these ideas we have accomplished a number of things. We first established that in a disordered system, a stable disordered Fermi–liquid FP is possible for $d > 2$. We showed that the so–called weak localization effects in itinerant electronic systems[48] are, or can be interpreted as, corrections to scaling near this FP. This derivation not only reproduces known perturbative results, but also establishes that their functional form is asymptotically exact in the long–wavelength limit. In this respect our achievement is analogous to that of Ref. 50 for long–time tails in classical fluids. In Section III B we have given a technically satisfactory derivation of the generalized nonlinear sigma–model that has been used to describe metal–insulator transitions near two dimensions. Finally, we have indicated how the theory must be modified near other quantum phase transitions where additional soft modes, namely the critical modes, appear.

For fermion systems without disorder, analogous results have been obtained. First, in agreement with others, we find that a clean Fermi–liquid FP exists for $d > 1$. Corrections to scaling near this FP show that clean electronic systems have nonanalyticities in various correlation functions that are analogous to weak–localization

effects in disordered systems. This is again in agreement with results obtained on the basis of either Fermi–liquid theory,[4] or many–body perturbation theory. Our treatment provides a unified description that reveals a deep connection between the behavior of fermionic systems with and without disorder, respectively. In Section IV A we also gave a novel perturbation theory method for clean fermion systems.

There are still a number of things to be done. For disordered systems, we recently suggested an order parameter description of the metal–insulator transition.[40] Our theory was based on the nonlinear sigma–model approach, applied to high dimensions (near $d = 6$), where its validity is not clear. This approach indicated that the metal–insulator transition has features in common with the transition in a classical random–field Ising model. These ideas need to be reexamined using a more general theory, since the sigma–model approach is asymptotically exact only for $d < 4$. Instanton solutions of the general field theory should also lead to insights concerning rare events, like local moment formation, and the effects of local moments on the metal–insulator transition. Finally, as already noted, additional quantum phase transitions from the Fermi–liquid state need to be investigated.

For clean systems, the most important thing to do is to include the effects of other Fermi–liquid parameters, as was discussed in Section II. Once this is done, quantum phase transitions in clean itinerant systems can be properly studied. It will then also be possible to study the crossover from the clean to the disordered Fermi liquid fixed point in detail. For non–interaction systems, this latter point has been addressed by Muzykantskii and Khmelnitskii.[61]

For both clean and disordered systems, it is interesting to ask how a Fermi–liquid FP can be avoided. This question has been of great interest recently in connection with high-T_c superconductors, and other systems that have 'strange metal' phases. For clean systems this has recently been reviewed in Ref. 62. For disordered systems, the situation is less clear. Various types of Kondo lattice mechanisms have been proposed.[63] In these approaches it is unclear how these effects modify the long–wavelength transport properties. Another possibility is to consider systems with a vanishing single–particle density of states at the Fermi surface. If this occurs, then the soft modes discussed in Sec. II are not as singular, and a two–dimensional disordered metal phase becomes possible. Further, it suggests the possibility of an exotic metal–insulator transition in two dimensions. Clearly, these problems require much more work.

for helpful discussions.

APPENDIX A: PROJECTION ONTO THE DENSITY

Here we demonstrate, for the spin–singlet particle–hole interaction term, the projection onto density modes that was used in Secs. II A and II B 1. An analogous procedure can be applied to the other interaction channels, and to the disorder part of the action.

It is most convenient to go back to a Hamiltonian description of the system. The part of the interaction Hamiltonian that corresponds to Eq. (2.10b) reads

$$H_{int}^{(s)} = \frac{1}{2} \sum_{\mathbf{k},\mathbf{p}} \sideset{}{'}\sum_{\mathbf{q}} \Gamma_{k,p}^{(s)}(q) \, \Delta f_{\mathbf{k}}(\mathbf{q}) \, \Delta f_{\mathbf{p}}(\mathbf{q}) \quad , \quad \text{(A1a)}$$

where

$$f_{\mathbf{k}}(\mathbf{q}) = \sum_{\sigma} c_{\sigma,\mathbf{k}}^{\dagger} \, c_{\sigma,\mathbf{k}+\mathbf{q}} \quad , \quad \text{(A1b)}$$

is the phase space density operator in terms of electron creation and annihilation operators $c^{\dagger}$ and c, respectively, and $\Delta(c^{\dagger}c) = c^{\dagger}c - \langle c^{\dagger}c \rangle$. In the space of products of fermion operators, we define the Kubo product,[64]

$$(A|B) = \int_0^{\beta} d\tau \, \langle \Delta B(\tau) \, \Delta A^* \rangle \quad , \quad \text{(A2)}$$

with A and B operators, and τ the imaginary time variable. In terms of this scalar product in operator space, the desired projector reads

$$\mathcal{P} = |n_n(\mathbf{q})) \frac{1}{g(\mathbf{q})} (n_n(\mathbf{q})| \quad . \quad \text{(A3)}$$

Here $n_n(\mathbf{q}) = \sum_{\mathbf{k}} f_{\mathbf{k}}(\mathbf{q})$ is the electron number density operator, and $g(\mathbf{q}) = (n_n(\mathbf{q})|n_n(\mathbf{q}))$ is the static density susceptibility or wave vector dependent compressibility. Using $\mathcal{P}$ twice, it is now easy to project onto the density in Eq. (A1a). We obtain

$$H_{int}^{(s)} \approx \frac{1}{2} \sideset{}{'}\sum_{\mathbf{q}} \Gamma^{(s)}(\mathbf{q}) \, n_n(\mathbf{q}) \, n_n(-\mathbf{q}) \quad , \quad \text{(A4a)}$$

where,

$$\Gamma^{(s)}(\mathbf{q}) = \frac{1}{g^2(\mathbf{q})} \sum_{\mathbf{k},\mathbf{p}} \Gamma_{k,p}^{(s)}(q) \, \big(f_{\mathbf{k}}(\mathbf{q})|n_n(\mathbf{q})\big)$$
$$\times \big(n_n(\mathbf{q})|f_{\mathbf{p}}(\mathbf{q})\big) \quad . \quad \text{(A4b)}$$

The phase space Kubo function, $g_{\mathbf{kp}}(\mathbf{q}) = (f_{\mathbf{k}}(\mathbf{q})|f_{\mathbf{p}}(\mathbf{q}))$, for clean, free electrons is proportional to $\delta_{\mathbf{kp}}\delta(\mathbf{k}^2 - k_F^2)$, so in this case Eq. (A4b) results in pinning $\mathbf{k}$ and $\mathbf{p}$ to the Fermi surface. In a disordered system, $g_{\mathbf{kp}}(\mathbf{q})$ has a width given by the inverse elastic mean–free path, and hence $\Gamma^{(s)}(\mathbf{q})$ is a weighted average over a region in

the vicinity of the Fermi surface, as mentioned in Sec. II A. Finally, for physics that is controlled by soft modes and long–wavelength processes we can replace $\Gamma^{(s)}(\mathbf{q})$ by $\Gamma^{(s)} \equiv \Gamma^{(s)}(\mathbf{q} = 0)$. Switching back to a field–theoretic representation of the fermions, we obtain Eq. (2.13a).

APPENDIX B: O(N) SYMMETRIC ϕ^4-THEORY

In this appendix we perform an analysis of ϕ^4-theory that is analogous to that of the matrix field theory in Sec. III B. Much of this material can be found in the literature, e.g. in Zinn-Justin's book, Ref. 35. However, we find it useful to include it here for pedagogical reasons. The remarkable analogy between field theories for electrons and classical spin models that was first noted by Wegner[14] substantially simplifies the understanding of the former in terms of the technically much simpler structure of the latter.

1. Origin of the nonlinear sigma–model

Let us consider an $O(N)$ symmetric ϕ^4-theory with a magnetic field h in the 1-direction. The action

$$
S[\vec{\phi}] = \int d\mathbf{x} \left[r\big(\vec{\phi}(\mathbf{x})\big)^2 + c\big(\nabla\vec{\phi}(\mathbf{x})\big)^2 \right]
$$
$$
+ u \int d\mathbf{x} \left(\vec{\phi}(\mathbf{x}) \cdot \vec{\phi}(\mathbf{x}) \right)^2 - h \int d\mathbf{x}\, \phi_1(\mathbf{x}) \quad , \quad \text{(B1a)}
$$

determines the partition function

$$
Z[h] = \int D[\vec{\phi}]\, e^{-S[\vec{\phi}]} \quad . \tag{B1b}
$$

In the low–temperature phase, where the $O(N)$ symmetry is spontaneously broken, it is convenient to decompose the vector field $\vec{\phi}$ into its modulus ρ and a unit vector field $\hat{\phi}$,

$$
\vec{\phi}(\mathbf{x}) = \rho(\mathbf{x})\,\hat{\phi}(\mathbf{x}) \quad , \quad \hat{\phi}^2(\mathbf{x}) \equiv 1 \quad , \tag{B2}
$$

$\hat{\phi}$ parametrizes the unit sphere, and thus provides a representation of the homogeneous space $O(N)/O(N-1)$. In terms of ρ and $\hat{\phi}$ the action reads,

$$
S[\rho, \hat{\phi}] = \int d\mathbf{x} \left[c^{(1)}\rho^2(\mathbf{x}) \left(\nabla\hat{\phi}(\mathbf{x}) \right)^2 + c^{(2)} \left(\nabla\rho(\mathbf{x}) \right)^2 \right.
$$
$$
\left. + r\rho^2(\mathbf{x}) \right] + u \int d\mathbf{x}\, \rho^4(\mathbf{x}) - h \int d\mathbf{x}\, \rho(\mathbf{x})\hat{\phi}_1(\mathbf{x}) ,
$$
$$
\tag{B3a}
$$

and switching from the functional integration variables $\vec{\phi}$ to $(\rho, \hat{\phi})$ leads to a Jacobian or invariant measure

$$
I[\rho] = \prod_{\mathbf{x}} \rho^{N-1}(\mathbf{x}) \quad . \tag{B3b}
$$

In Eq. (B3a) the bare values of $c^{(1)}$ and $c^{(2)}$ are equal, and equal to c. Notice that the field $\hat{\phi}$ appears only in conjunction with two gradient operators. $\hat{\phi}$ represents the $N-1$ soft Goldstone modes of the problem, while ρ represents the massive modes. Now we parametrize $\hat{\phi}$,

$$
\hat{\phi}(\mathbf{x}) = \big(\sigma(\mathbf{x}), \vec{\pi}(\mathbf{x})\big) \quad , \tag{B4a}
$$

where

$$
\sigma(\mathbf{x}) = \sqrt{1 - \vec{\pi}^2(\mathbf{x})} \quad . \tag{B4b}
$$

We split off the expectation value of the massive ρ-field, $\rho(\mathbf{x}) = M + \Delta\rho(\mathbf{x})$, with $M = \langle\rho(\mathbf{x})\rangle$, and expand in powers of $\vec{\pi}$ and $\Delta\rho$. Rescaling the coupling constants with appropriate powers of M, the action can be written

$$
S[\rho, \vec{\pi}] = S_{NL\sigma M}[\vec{\pi}] + \Delta S[\rho, \vec{\pi}] \quad . \tag{B5a}
$$

Here

$$
S_{NL\sigma M}[\vec{\pi}] = \frac{1}{t} \int d\mathbf{x} \left[(\nabla\vec{\pi}(\mathbf{x}))^2 + (\nabla\sigma(\mathbf{x}))^2 \right]
$$
$$
- h \int d\mathbf{x}\, \sigma(\mathbf{x}) \quad , \tag{B5b}
$$

is the action of the $O(N)$ nonlinear sigma–model, and

$$
\Delta S[\rho, \vec{\pi}] = r \int d\mathbf{x}\, (\Delta\rho(\mathbf{x}))^2 + c^{(2)} \int d\mathbf{x}\, (\nabla\Delta\rho(\mathbf{x}))^2
$$
$$
+ u_3 \int d\mathbf{x}\, (\Delta\rho(\mathbf{x}))^3 + u_4 \int d\mathbf{x}\, (\Delta\rho(\mathbf{x}))^4
$$
$$
+ O\left(\Delta\rho\,\sigma, \Delta\rho \left(\nabla\hat{\phi}\right)^2 \right) \quad . \tag{B5c}
$$

If we neglect all fluctuations of the ρ-field, then we are left with the $O(N)$ symmetric nonlinear sigma–model in the usual parametrization.

2. The low–temperature fixed point

Now we define the scale dimensions of the fields π_i and $\Delta\rho$ as

$$
[\pi_i(\mathbf{x})] = \frac{1}{2}(d - 2 + \eta') \quad , \tag{B6a}
$$

$$
[\Delta\rho(\mathbf{x})] = \frac{1}{2}(d - 2 + \eta) \quad , \tag{B6b}
$$

and perform a momentum–shell RG procedure.[41] Here we have defined the scale dimension of a length L to be $[L] = -1$, and the above relations define the exponents η and η'. For η, this definition coincides with that of the exponent usually denoted by this symbol. The stable, low–temperature FP of the action describes the ordered phase. Physically, one expects short–ranged

$\Delta\rho$–correlations, and diffusive π–correlations at this FP, which suggests the choice[65]

$$\eta = 2 \quad , \quad \eta' = 0 \quad . \tag{B7}$$

Because of the positive scale dimension of $\vec{\pi}$, it is convenient to expand the sigma–model action,

$$S_{NL\sigma M} = \frac{1}{t} \int d\mathbf{x} \, (\nabla\vec{\pi}(\mathbf{x}))^2 + \frac{h}{2} \int d\mathbf{x} \, \vec{\pi}^2(\mathbf{x})$$
$$+ v \int d\mathbf{x} \, (\vec{\pi}(\mathbf{x})\nabla\vec{\pi}(\mathbf{x}))^2 + O\left(\vec{\pi}^4, \nabla^2\vec{\pi}^6\right) \quad . \tag{B8}$$

The FP action is given by the first terms in Eqs. (B8) and (B5c), respectively. With respect to this FP, the magnetic field is the only relevant operator, with a scale dimension $[h] = 2$. t and r are marginal, and all other operators are irrelevant with respect to this FP, so the FP is really stable.

The leading correction to scaling near $d = 2$ is given by the operator v, with a scale dimension $[v] = -(d-2)$. This identifies $d_c^- = 2$ as the lower critical dimension of the problem. v also provides the leading nonanalytic correction to scaling in *all* dimensions, since the leading irrelevant operator related to the massive $\Delta\rho$-field has a scale dimension $[c^{(2)}] = -2$, and therefore leads to analytic corrections to scaling. For $d > d_c^-$, the spontaneously broken $O(N)$ symmetry leads to $N-1$ Goldstone modes, which are represented by the correlations of the $n - 1$ π-fields,

$$\langle \pi_i(\mathbf{k}) \, \pi_j(-\mathbf{k}) \rangle = \delta_{ij} |\mathbf{k}|^{-2+\eta'} = \delta_{ij}/\mathbf{k}^2 \quad . \tag{B9}$$

The FP value of the magnetization, $m = \langle \phi_1(\mathbf{x}) \rangle$, is equal to M, and the leading correction is given by the π-π correlation function. By virtue of the scale dimensions of π and h, one finds for the leading field dependence of the magnetization

$$m(h) = \text{const.} + h^{(d-2)/2} \quad , \tag{B10}$$

which implies that the longitudinal susceptibility, $\chi_L = \partial m/\partial h \sim h^{(d-4)/2}$, diverges in the $h \to 0$ limit for all $d < 4$.[66] Alternatively, the wave vector dependent zero–field susceptibility diverges in the homogeneous limit like $\chi_L \sim |\mathbf{k}|^{d-4}$. Notice that the leading nonanalytic corrections to scaling are given by the Goldstone modes, and are thus contained in the nonlinear sigma–model.

The point of the above excercise, in the present context, is to demonstrate how much more one gets out of power counting after performing a symmetry analysis and separating the soft modes from the massive ones, as opposed to doing the power counting on the action in the original $\vec{\phi}$ formulation.[67] The analogy to the procedure in Sec. III B is as follows: $\Delta\rho$ corresponds to the massive fields ΔP and $\Delta\Lambda$, $\hat{\phi}$ corresponds to $\hat{Q}$ and $\mathcal{S}$ (the matrix theory cannot be expressed in terms of $\hat{Q}$ only), and $\vec{\pi}$

corresponds to q. The external magnetic field h plays the role of the external frequency in the matrix theory, and the analogy with respect to the Goldstone modes and the scaling behavior of the order parameter is obvious.

3. The critical fixed point

Apart from the low–temperature FP discussed above, the nonlinear sigma–model, Eq. (B5b), also contains a critical FP, as was noticed by Polyakov,[68] and discussed by him and others.[69–71]. This FP marks the instability of the low–temperature or broken symmetry phase; the $O(N)$ symmetry is restored, and hence the fields $\vec{\pi}$ and σ have the same scale dimension. $\vec{\pi}$ must then be dimensionless, which implies

$$\eta' = 2 - d \quad , \tag{B11a}$$

in Eq. (B6a). The exponent η in Eq. (B6b), on the other hand, is the critical correlation function exponent that is related to the order parameter exponent β, and the susceptibility exponent γ, via

$$\eta = 2 - \gamma/\nu = 2 - d + 2\beta/\nu \quad . \tag{B11b}$$

At criticality, the $\Delta\rho$ fluctuations become massless, and the region of validity of the sigma–model shrinks to zero. Specifically, the $\Delta\rho$ fluctuations can be neglected only if one works at momenta p that are larger then the mass of the $\Delta\rho$-field, $m_\rho = \sqrt{r/c^{(2)}}$, see Eq. (B5c). Power counting, and Eq. (B6b), shows that the scale dimension of m_ρ is $[m_\rho] = 1$, so upon approaching criticality it vanishes like $m_\rho \sim \xi^{-1} \sim |T - T_c|^\nu$. The criterion for the validity of the sigma–model is therefore

$$p \ll |T - T_c|^\nu \quad . \tag{B12}$$

In this region perturbation theory works, and the critical properties can be explored by using the RG to sum the perturbative results. In this way one obtains a description of the Heisenberg critical behavior in the vicinity of $d = 2$ that complements the perturbative treatment of the ϕ^4-formulation of the problem, Eq. (B1a), near $d = 4$. The critical FP of the matrix nonlinear sigma–model, Sec. III B 3, is again in many respects analogous to the critical FP of the $O(N)$ model.

It has been suggested that terms with more than two gradients, which appear as corrections to the sigma–model upon explicitly integrating out the massive field, may lead to an instability of the critical FP, even though these terms are irrelevant by power counting at tree level.[72] The status of this problem is currently unclear, see Ref. 73 for a recent discussion.

APPENDIX C: THE INVARIANT MEASURE $I[P]$

Here we derive explicitly the invariant measure $I[P]$, defined by Eq. (3.25). Let us write Eq. (3.24) in block matrix form,

$$
\begin{pmatrix} Q^{>>} & Q^{><} \\ Q^{<>} & Q^{<<} \end{pmatrix} = \begin{pmatrix} \mathcal{S}^{>>} & \mathcal{S}^{><} \\ \mathcal{S}^{<>} & \mathcal{S}^{<<} \end{pmatrix} \begin{pmatrix} P^{>} & 0 \\ 0 & P^{<} \end{pmatrix}
$$
$$
\times \begin{pmatrix} (\mathcal{S}^{-1})^{>>} & (\mathcal{S}^{-1})^{><} \\ (\mathcal{S}^{-1})^{<>} & (\mathcal{S}^{-1})^{<<} \end{pmatrix} \quad . \quad \text{(C1)}
$$

Here each block is a $4Nn \times 4Nn$ matrix. P is block diagonal by construction, and the blocks of $\mathcal{S}$ are subject to constraints due to the definition of the coset space. Now consider a variation δQ of Q. Since by symmetry the invariant measure is independent of the group parameters, it suffices to consider an infinitesimal variation. To first order in infinitesimal quantities, the variation of $\mathcal{S}$ is block off–diagonal (see Eq. (3.48)), and we have,

$$
\begin{pmatrix} \delta Q^{>>} & \delta Q^{><} \\ \delta Q^{<>} & \delta Q^{<<} \end{pmatrix} =
$$
$$
\begin{pmatrix} \delta P^{>} & \delta \mathcal{S}^{><} P^{<} - P^{>} \delta \mathcal{S}^{><} \\ \delta \mathcal{S}^{<>} P^{>} - P^{<} \delta \mathcal{S}^{<>} & \delta P^{<} \end{pmatrix} \quad .
$$
$$\text{(C2)}$$

By directly differentiating Q with respect to P we find from Eq. (C2),

$$
I[P] = \det \left[\mathbb{1} \otimes (P^{<})^{T} - P^{>} \otimes \mathbb{1} \right] \quad . \quad \text{(C3)}
$$

$P^{>}$ and $P^{<}$ are analytic continuations of anti–hermitian matrices, so they can be diagonalized by means of analytic continuations of unitary $4Nn \times 4Nn$ matrices. Let their eigenvalues be $\lambda_i^{>}$ and $\lambda_i^{<}$ ($i = 1, \ldots 4Nn$). Putting the space dependence back in, we finally obtain

$$
I[P] = \left| \prod_{\mathbf{x}} \prod_{i,j=1}^{4Nn} \left(\lambda_i^{<}(\mathbf{x}) - \lambda_j^{>}(\mathbf{x}) \right) \right| \quad . \quad \text{(C4)}
$$

[1] D. Pines and P. Nozières, *The Theory of Quantum Liquids*, vol.1, Addison–Wesley (New York 1989).

[2] A. L. Fetter and D. L. Walecka, *Quantum Theory of Many–Particle Systems*, Mc Graw–Hill (New York 1971).

[3] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics*, Dover (New York 1975).

[4] See, e.g., G. Baym and C. Pethick, *Landau Fermi–Liquid Theory*, John Wiley (New York 1991).

[5] C. Castellani and C. Di Castro, Phys. Rev. B **34**, 5935 (1986).

[6] C. Castellani, G. Kotliar, and P. A. Lee, Phys. Rev. Lett. **59**, 323 (1987).

[7] See, e.g., J.W. Negele and H. Orland, *Quantum Many–Particle Systems*, Addison Wesley (New York 1988).

[8] R. Shankar, Rev. Mod. Phys. **66**, 129 (1994).

[9] N. Dupuis and G. Y. Chitov, Phys. Rev. B **54**, 3040 (1996); N. Dupuis, cond-mat/9604189.

[10] A. Houghton and J. B. Marston, Phys. Rev. B **48**, 7790 (1993); A. Houghton, H. J. Kwon, and J. B. Marston, Phys. Rev. B **50**, 1351 (1994); H. J. Kwon, A. Houghton, and J. B. Marston, Phys. Rev. B **52**, 8002 (1995); A. Houghton, H. J. Kwon, J. B. Marston, and R. Shankar, J. Phys. Cond. Matter **6**, 4909 (1994).

[11] J. Fröhlich and R. Götschmann, Phys. Rev. B **55**, 6788 (1997).

[12] P. Kopietz, J. Hermisson, and K. Schönhammer, Phys. Rev. B **52**, 10877 (1995); P. Kopietz and K. Schönhammer, Z. Phys. B **100**, 259 (1996).

[13] P. Kopietz, J. Phys.: Condens. Matter **8**, 10483 (1996).

[14] F. Wegner, Z. Phys. B **35**, 207 (1979).

[15] A. M. Finkel'stein, Zh. Eksp. Teor. Fiz. **84**, 168 (1983) [Sov. Phys. JETP **57**, 97 (1983)]; Z. Phys. B **56**, 189 (1984).

[16] For a review, see, D. Belitz and T. R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).

[17] L. Schäfer and F. Wegner, Z. Phys. B **38**, 113 (1980). See also A. J. McKane and M. Stone, Ann. Phys. **131**, 36 (1981).

[18] A. M. M. Pruisken and L. Schäfer, Nucl. Phys. B **200[FS4]**, 20 (1982).

[19] T. R. Kirkpatrick and D. Belitz, Phys. Rev. B **53**, 14364 (1996).

[20] T. Vojta, D. Belitz, R. Narayanan, and T. R. Kirkpatrick, Europhys. Lett. **36**, 191 (1996); cond-mat/9611099 (Z. Phys. B **xx**, xxx (1997)).

[21] T. R. Kirkpatrick and D. Belitz, Phys. Rev. Lett. **76**, 2571 (1996) (Erratum ibid. **78**, 1197 (1997)).

[22] For a pedagogical discussion of the replica trick, see, e.g., G. Grinstein, in *Fundamental Problems in Statistical Mechanics VI*, edited by E. G. D. Cohen, North Holland (Amsterdam 1985), p.147.

[23] To avoid possible misconceptions, we mention that this is unrelated to the fact that n_n and $\mathbf{n}_s$ are conserved quantities. The softness will turn out to be due to a spontaneously broken continuous symmetry, *not* due to a conservation law. Indeed, n_c is not a conserved quantity. See also the discussion after Eqs. (3.19).

[24] K. B. Efetov, A. I. Larkin, and D. E. Khmelnitskii, Zh. Eksp. Teor. Fiz. **79**, 1120 (1980) [Sov. Phys. JETP **52**, 568 (1980)].

[25] Note that the adjoint operation defined here has the properties $(\eta^{+})^{+} = \eta$, and $(z\eta)^{+} = z\eta^{+}$, with $z \in \mathcal{C}$. It is unrelated to the adjoint operation defined for the ψ, which obeys $\overline{\overline{\psi}} = \psi$, and $\overline{z\psi} = z^{*}\overline{\psi}$. Both our notation and our definition deviates slightly from Refs. 24 and 16 at this point.

[26] Instead of this contraint procedure, which introduces the 'ghost' field $\tilde{\Lambda}$, one might think of using a Hubbard–Stratonovich decoupling of the four–fermion terms. However, the interaction terms are not easily decoupled, since the corresponding operators have zero eigenvalues.[16] The disorder term can be easily decoupled, but after doing so the clean limit is highly singular, and not easily performed. We therefore choose the ghost–field method, which does not lead to any of these problems, and allows for the clean limit to be taken explicitly at any point.

[27] T. R. Kirkpatrick and D. Belitz, unpublished.

[28] See, e.g., N. F. Mott, *Metal-Insulator Transitions*, Taylor&

Francis (London 1990), and references therein.

[29] See, Sec. II.A.1.d in Ref. 16, and references therein.

[30] The reason why the single–particle relaxation time τ_{rel}, rather than the transport time τ_{tr}, appears in the diffusion coefficient lies in our projecting on the density only in Sec. II B 1. This is sufficient to correctly yield the cancellation between the leading terms in the brackets in Eq. (2.40a), but not for giving the correct prefactor of the next–to–leading term that determines D. If desired, this can be easily remedied by keeping the longitudinal current mode in the projector, Eq. (A3). For our purposes, however, the distinction between τ_{rel} and τ_{tr} is irrelevant.

[31] B. G. Wybourne, *Classical Groups for Physicists*, Wiley (New York 1974); R. Gilmore, *Lie Groups, Lie Algebras, and Some of Their Applications*, Wiley (New York 1974). We will not always distinguish between the abstract groups and their matrix representations, since this distinction is not necessary for our purposes.

[32] This identity is a simple extension of the relation $\langle G \rangle = -2i\langle Q \rangle$, see Ref. 18, and can be proven in the same way.

[33] S. V. Maleev and B. P. Toperverg, Zh. Eksp. Teor. Fiz. **69**, 1440 (1975) [Sov. Phys. JETP **42**, 734 (1976)].

[34] Since all of the replicas are identical, quantities that depend only on one replica index are replica independent.

[35] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Clarendon (Oxford 1989), ch. 27.

[36] Strictly speaking, the $Q^{\alpha\beta}_{nm,\sigma\sigma'}$ are not elements of the usual quaternion field (defined as the set of $q = \sum_{r=0}^{3} q_r \tau_r$ with $q_r \in \mathcal{R}$), since they are not real with respect to the quaternion basis. However, the set of Q obeying Eqs. (2.26) is isomorphic to $\mathcal{Q}$, since these properties are invariant under quaternion multiplication.

[37] Here we use the fact that if T is symplectic, then so is T^T, which can be seen as follows. Let T be symplectic, so $T^T C T = C$. Then $T^{-1} = (C^T T C)^T$, since $C^T = C^{-1}$. But $(T^T)^{-1} = (T^{-1})^T$ for all T, and hence $(T^T)^{-1} = C^T T C$, or $T C T^T = C$.

[38] Notice that the fluctuations of $\tilde{\Lambda}$ are *not* massive, cf. Eqs. (2.37), (2.38). The transformation from $\tilde{\Lambda}$ to Λ here has the same effect as switching from $\tilde{\Lambda}$ to $\bar{\Lambda}$ in the Gaussian approximation, Sec. II C, namely it subtracts the trivially soft components of the ghost field that result from its coupling to Q.

[39] This can also be shown directly by defining the Green operator, $\hat{G}$, whose matrix elements form the Green function G, and considering matrix elements of the commutator $[\hat{x}, \hat{G}^{-1}]$. See also Ref. 18.

[40] T. R. Kirkpatrick and D. Belitz, Phys. Rev. Lett. **73**, 862 (1994); ibid. **74**, 1178 (1995); D. Belitz and T. R. Kirkpatrick, Z. Phys. B **98**, 513 (1995).

[41] K. G. Wilson and J. Kogut, Phys. Rep. **12**, 75 (1974).

[42] Instead, one could use a momentum-frequency shell RG (see, J. Hertz, Phys. Rev. B **14**, 1165 (1976)). For the problem under consideration, this would lead to a coupling between the large and small frequency degrees of freedom even at the Gaussian level. Since we will conclude that performing the RG transformation on the Gaussian theory is sufficient, this is an unnecessary complication. However, if performed exactly, this more complicated procedure is equivalent to the simpler one we choose to use.

[43] We are using Ma's method of identifying simple fixed points. That is, we fix the scale dimensions of the fields, as well as the dynamical scaling exponent, and then self–consistently verify that this choice leads to a stable FP. See S.-K. Ma, *Modern Theory of Critical Phenomena*, Benjamin (Reading, MA 1976), Sec. VII.1.

[44] L. P. Gorkov, A. I. Larkin, and D. E. Khmelnitskii, Pis'ma Zh. Teor. Fiz. **30**, 248 (1979) [JETP Lett. **30**, 228 (1979)].

[45] B. L. Altshuler, A. G. Aronov, D. E. Khmelnitskii, and A. I. Larkin, in *Quantum Theory of Solids*, edited by I. M. Lifshits, Mir Publishers (Moscow, 1982), p.130; B. L. Altshuler and A. G. Aronov, in *Electron-Electron Interactions in Disordered Systems*, A. L. Efros and M. Pollak, eds., North Holland (Amsterdam 1985), p.1.

[46] C. Castellani, C. Di Castro, G. Kotliar, P. A. Lee, and G. Strinati, Phys. Rev. B **37**, 9046 (1988).

[47] B. L. Altshuler and A. G. Aronov, Solid State Commun. **30**, 115 (1979).

[48] The term 'weak localization' is ill–defined, and often misunderstood. Here we use it to refer to the nonanalytic behavior of electronic correlation functions in the limit of zero momentum and/or frequency that is induced by quenched disorder, or by a combination of interactions and quenched disorder, and that occurs even if the system is far from any kind of phase transition. The physical reason for these nonanalyticities is the diffusive motion of the electrons in the presence of quenched disorder. For a detailed discussion see, e.g., Ref. 45.

[49] For a review, see, e.g., J. R. Dorfman, T. R. Kirkpatrick, and J. V. Sengers, Ann. Rev. Phys. Chem. **45**, 213 (1994). See also T. R. Kirkpatrick and D. Belitz, J. Stat. Phys. **87**, 1307 (1997).

[50] D. Forster, D. R. Nelson, and M. J. Stephen, Phys. Rev. A **16**, 732 (1977).

[51] This is true only for interacting electrons, when the metal–insulator transition is of the Anderson–Mott type (and even then not for all of the possible universality classes, see Ref. 16). For noninteracting electrons, one has an Anderson transition, which is *not* of the symmetry restoring type, as has been pointed out by McKane and Stone, Ref. 17.

[52] Notice that part of the scale dimension of Q has been lumped into the coupling constant G, see Ref. 53. We also note that at the critical FP, the numerical value of η cannot be as easily identified as at the stable Fermi–liquid FP. In order to obtain numerical values one has to employ some kind of ϵ-expansion scheme.

[53] This is true despite the fact that the scale dimension of q is zero at the critical FP according to Eqs. (3.49a), (3.65). The salient point is that in going from the theory in terms of Q, Eq. (3.24), to that in terms of $\hat{Q}$, Eqs. (3.40b), (3.47), the scale dimension of P has been absorbed in the coupling constant G in Eq. (3.51).

[54] D. Belitz and T. R. Kirkpatrick, Phys. Rev. B **52**, 13922 (1995).

[55] D. Belitz and T. R. Kirkpatrick, Phys. Rev. B **44**, 955 (1991). T. R. Kirkpatrick and D. Belitz, Phys. Rev. B **45**, 3187 (1992).

[56] See, e.g., H. J. Schulz *Fermi Liquids and Non–Fermi Liquids*, in "Proceedings of Les Houches Summers School

LXI", E. Akkermans, G. Montambaux, J. Pichard, and J. Zinn–Justin, eds., Elsevier (Amsterdam, 1995), p.533 (cond-mat/9503150).

[57] P. W. Anderson, Phys. Rev. Lett. **64**, 1839 (1990); *ibid.* **65**, 2306 (1990). See also P. C. E. Stamp, Phys. Rev. Lett. **68**, 2180 (1992), and P. W. Anderson and D. Khveshchenko, Phys. Rev. B **52**, 16415 (1996).

[58] G. Benfatto and G. Gallavotti, J. Stat. Phys. **59**, 541 (1990); Phys. Rev. B **42**, 9967 (1990).

[59] H. Fukuyama, O. Nariyiko, and Y. Hasegawa, J. Phys. Soc. Jpn. **60**, 372 (1991); J. R. Engelbrecht and M. Randeria, Phys. Rev. B **45**, 12419 (1992); C. Castellani, C. Di Castro, and W. Metzner, Phys. Rev. Lett. **72**, 316 (1994).

[60] D. Belitz, T.R. Kirkpatrick, and T. Vojta, Phys. Rev. B **55**, 9452 (1997).

[61] B. A. Muzykantskii and D. E. Khmelnitskii, Pis'ma Zh. Eksp. Teor. Fiz. **62**, 68 (1995).

[62] W. Metzner, C. Castellani, and C. Di Castro, cond-mat/9701012.

[63] See, e.g., *Institute for Theoretical Physics Conference on Non–Fermi Liquid Behavior in Metals*, P. Coleman, B. Maple, and A. Millis (eds.), J. Phys. Cond. Matt. **48**, No. 8 (1996).

[64] R. Kubo, J. Phys. Soc. Japan **12**, 570 (1957).

[65] This choice for the scale dimensions of the fields is not unique. It is, however, particularly convenient for a discussion of the FP in question. Other choices for η and η' lead to the appearance of dangerous irrelevant variables, which makes the discussion of the scaling behavior near the FP much more cumbersome.

[66] E. Brézin and Wallace, Phys. Rev. B **7**, 1967 (1973).

[67] S.-K. Ma, Ref. 43.

[68] A. M. Polyakov, Phys. Lett. **59**, 79 (1975).

[69] E. Brézin and J. Zinn-Justin, Phys. Rev. B **14**, 3110 (1976).

[70] E. Brézin, J. Zinn-Justin, and J. C. Le Guillou, Phys. Rev. D **14**, 2615 (1976).

[71] D. R. Nelson and R. A. Pelcovits, Phys. Rev. B **16**, 2191 (1977).

[72] V. E. Kravtsov, I. V. Lerner, and V. I. Yudson, Phys. Lett. A **134**, 245 (1989); F. Wegner, Z. Phys. B **78**, 33 (1990); G. E. Castilla and S. Chakravarty, Nucl. Phys. B **485**, 613 (1997).

[73] E. Brézin and S. Hikami, cond-mat/9612016.